Acc. No   ... .

ISSUE LABEL

Not later than the latest date stamped below.

COLLOID SCIENCE

II

COLLOID SCIENCE

Edited by

H. R. KRUYT

Formerly Professor of Physical Chemistry, Utrecht University
President of the Central National Council
for Applied Scientific Research in the Netherlands, T.N.O., The Hague

VOLUME II

REVERSIBLE SYSTEMS

ELSEVIER PUBLISHING COMPANY, INC.

NEW YORK - AMSTERDAM - LONDON - BRUSSELS

1949

ENGLISH TRANSLATIONS BY
DR. L C. JACKSON (BRISTOL)

PREFACE

The need for a collective work on the science of colloids formed a subject of discussion between the publishers of this work and myself shortly before the outbreak of the Second World War. The first preparations were brought to an end by the world events, in view of the fact that all possibility of contact with foreign colleagues had been broken. When the war continued year after year, the question arose whether we could restrict ourselves to Dutch collaborators and in 1944 I sent a request to a number of my fellow-countrymen to take this work in hand. It was immediately apparent that this plan was received with enthusiasm and most of those who now collaborate in this work met to devise a plan of campaign; a few others, whom we invited to do so, joined us later.

In our first discussion two points especially occupied our attention. In the first place the question was asked whether "colloid science" is an adequately definable concept. In the second place the question arose whether colloid science is a sufficiently homogeneous domain so that the work could be knit together by a general line of thought and would not degenerate into a collection of separate communications without inner coherence.

We believed that both questions can be answered in the affirmative. To be sure each classification of science has always a restricted significance, because it ruptures general connections more or less arbitrarily but the method of human thought does call for classification. Colloidal systems are characterised by their possession of large kinetic units; there is undoubtedly continuity between the behaviour of large and small kinetic units but there are sufficiently salient differences between the extremes to divide the treatment without, however, losing sight of the connections.

We believe that we must also answer in the affirmative the question whether sufficient harmony exists within the domain of systems with large kinetic units and we are of the opinion that it serves a useful purpose just now to try to accentuate that harmony. Many modern books in the domain of colloid science have a monographic character and the impression is created that there is no connection between the "classical" colloid science as typified by FREUNDLICH's *Colloid and Capillary Chemistry* and the whole domain of macromolecular compounds which has developed so greatly in the last two decennia.

Nevertheless there is that harmony although the behaviour of the systems with large kinetic units is so diverse that one is inclined to lose sight of the connection. We are of the opinion that the following classification must be kept in mind.

In the first place one should distinguish reversible and irreversible systems, that is to say, colloid systems which can undergo phase change or phase separation reversibly or otherwise. A thermodynamically definable stability difference is thus the basis of this classification. It is for this reason more logical than the old classification into lyophobic and lyophilic systems (and to a still greater degree than that into suspensoids and emulsoids).

Within each of the two groups a second distinction can be made, namely whether one is dealing with systems with or without electrolytic character, taking this term in the broad sense. The irreversible systems have an electrolytic character whereby

the electric double layer plays a great part, as well as the protein systems which for the rest belong to the reversible group. The synthesized macromolecular substances will often belong to the reversible systems without electrolytic character.

With all the diversity which is recognised on the basis of this classification, there remains an enormous number of properties which these systems have in common and which are just connected with the fact that they are built up from large kinetic units.

For practical reasons it appeared desirable to divide this work into two volumes; the main division is this, that Volume I, after a general introduction, deals with the *irreversible systems*. This Volume consists of the following chapters;

    I.   General Introduction. H. R. KRUYT with the collaboration of J. J. HERMANS.
   II. Phenomenology of hydrophobic systems. J. TH. G. OVERBEEK.
  III. Optical properties of colloidal systems. G. H. JONKER.
  IV. Electrochemistry of the double layer. E. J. W. VERWEY and J. TH. G. OVERBEEK.
   V. Electrokinetic phenomena. J. TH. G. OVERBEEK.
  VI. The interaction between colloidal particles. J. TH. G. OVERBEEK and E. J. W. VERWEY.
 VII. Kinetics of flocculation. J. TH. G. OVERBEEK.
VIII. Stability of hydrophobic colloids and emulsions. J. TH. G. OVERBEEK and E. J. W. VERWEY.
  IX. Rheology. G. H. JONKER.
   X. Miscellaneous subjects. G. H. JONKER, J. TH. G. OVERBEEK and E. J. W. VERWEY.

Volume II, Macromolecular and Association Colloids, deals with the *reversible systems*. For the further subdivision of Volume II, reference may be made to the Contents and to Chapter I, § 5, especially, p. 16.

Through accidental circumstances Volume II has been completed earlier than Volume I and thus is published first.

Anyone who examines the outline of Volume I and the Contents of Volume II will decide that this work does not cover the whole of colloid science. This may readily be conceded and was done deliberately. We have been conscious from the beginning that a sufficiently expert collaborator was not to be found within our circle for every branch; I may mention the disperse systems in gaseous media as an example. In such cases we have deliberately waived the treatment of them. In addition there are branches about which excellent monographs exist; as an example I may mention the application of the ultracentrifuge. In such cases we arranged to treat the subject only in so far as it demanded discussion in a general way and for the rest to refer to these monographs.

The division of the treatment is made throughout according to general phenomena. The examples are taken each time from those systems in which the phenomenon is shown most characteristically or has been studied most satisfactorily. A systematic treatment of the properties of particular substances or systems (for example, of the proteins) is not to be expected in these volumes.

It will strike the reader that in some chapters, in particular those written by Prof. BUNGENBERG DE JONG, the treatment is relatively ample and many experimental data are given. We intended in this way to meet the objection that much of the original literature on this branch is less accessible to English and American readers.

The present work has no pretention of being a complete treatise. It is only meant to be a guide to the domain of colloid science with the subject of providing a stimulus in the branch of research with which it deals.

It is a pleasure to express my gratitude to those who have written the greater part of this work and with whom it has been a joy to be associated. I must however express a separate word of thanks to Prof. OVERBEEK who, when I had left the University of Utrecht and was too much occupied with other work, took upon himself the actual editing of this work.

The Hague, January 1949.   H. R. KRUYT.

COLLABORATORS

H. R. KRUYT

Formerly Professor of Physical Chemistry, Utrecht University
President of the Central National Council for Applied Scientific Research in the Netherlands, T.N.O., The Hague

H. L. BOOY

Lecturer of Medical Chemistry, Leyden University

H. G. BUNGENBERG DE JONG

Professor of Medical Chemistry, Leyden University

J. J. HERMANS

Professor of Physical Chemistry, Groningen University

P. H. HERMANS

Director of the Institute for Cellulose Research of the A.K.U. and affiliated companies, Utrecht

R. HOUWINK

General Director of the Rubber Foundation, Delft
External Lecturer, Delft Technical University

J. TH. G. OVERBEEK

Professor of Physical Chemistry, Utrecht University

CONTENTS

III. THERMODYNAMICS OF LONG-CHAIN MOLECULES

by J. J. Hermans

IV. THE PHYSICAL PROPERTIES OF RANDOMLY KINKED LONG CHAIN MOLECULES

by J. J. Hermans

IX. REVERSAL OF CHARGE PHENOMENA, EQUIVALENT WEIGHT AND SPECIFIC PROPERTIES OF THE IONISED GROUPS

by H. G. Bungenberg de Jong

X. COMPLEX COLLOID SYSTEMS

by H. G. Bungenberg de Jong

XI. MORPHOLOGY OF COACERVATES

by H. G. Bungenberg de Jong

XII. GELS
by P. H. Hermans

XIII. SOLID MACROMOLECULAR SYSTEMS WITH ONE (CHEMICAL) COMPONENT

by R. Houwink

XIV. ASSOCIATION COLLOIDS

by H. L. Booij

I. A SURVEY OF THE STUDY OBJECTS IN THIS VOLUME

H. G. BUNGENBERG DE JONG

Department of Medical Chemistry, University of Leyden

§ 1. DEFINITIONS

a. The historical development of two lines of thought

Readers not in possession of Volume I will·certainly be interested in the meaning of the word *Colloid Science* and a few other terms in frequent use. Before giving their definitions a short and far from complete historical survey may be included, from which the basic principles, on which the definitions are based may be recognized.

GRAHAM introduced the term *"Colloids"* for substances such as gelatin and other proteins, which in "solution" showed only a very slow diffusion velocity as compared with others such as sugar and salts. He did so because he considered gelatin ($\dot{\eta}$ $\varkappa o\lambda\lambda a$ = glue) as a representative example.

At that time ideas about the nature of solutions were still somewhat confused. As they cleared up later, it was recognized that the abnormal behaviour of these "colloids" — showing also abnormally low osmotic pressure, low values for the depression of the freezing point and increase of the boiling point — may have two different causes:

a. the dissolved substance is in true solution, but has a very high molecular weight,

b. the substance is present in a very fine but not yet molecular subdivision, forming a *"Pseudo-solution"* or *"Sol"* with the liquid in which the still polymolecular particles are suspended.

Explanation *a*. was frequently preferred by workers in the field of physiological chemistry for their special study objects such as proteins and other substances of biological origin.

Physical chemists, with their natural fear of working with ill-defined substances were however specially interested in colloidal solutions, which resulted from known chemical reactions and in which the substance present was also a known chemical substance.

As a result of the intensive study of such colloidal solutions, as those of gold, As_2S_3 and others, hardly any doubt could remain about the validity of explanation *b*. The polymolecularity of their kinetic units was definitely settled as true after the discovery of the ultra-microscope (ZSIGMONDY).

Following these. successes an overestimation of the value of explanation *b*. set in, it being believed valid in all cases. The term "Colloid" as a class name in GRAHAM's sense of the word became obsolete and *"Colloid Chemistry"* might subsequently be defined as the physical chemistry of two-phase systems, one of the phases being dispersed to so-called colloidal dimensions in the bosom of the other phase.

In this trend of thought an extensive knowledge of the properties of phase boundaries seemed of utmost importance for all study objects of colloid chemistry. So FREUNDLICH in his well known textbook (1930) — significantly entitled "Kapillarchemie" — could write in his general introduction, that colloid chemistry must be founded on three pillars viz., the knowledge of phase boundary phenomena, that of Brownian movement, and that relating to the laws governing the formation of new phases. Meanwhile workers in the field of protein solutions, in most cases still preferring explanation a. for their study objects, developed their methods more and more and were able to bring forward much evidence in favour of a. After the results obtained with the ultracentrifuge (THE SVEDBERG) there can be no longer any doubt that many proteins are molecularly dissolved in solution.

b. Definitions

As shown above, "Colloid Chemistry" has to deal with different "States of Solution" and its system of basic definitions should be founded therefore on the only link uniting them, that is, on the great mass of their kinetic units.

Using this principle we give the following definitions:

1. *Colloid Science*, the science, which deals a. with kinetic units of relatively great mass, the latter lying between arbitrarily chosen limits, b. with material systems containing these kinetic units or structures derived from them.

2. *Colloidal Particles*, the above named kinetic units themselves.

3. *Colloidal System*, a (macro)-homogeneous material system containing colloidal particles or structures derived from them, the presence of smaller kinetics units, falling outside the definition of colloidal particles, not being excluded.

4. *Sol*, a colloidal system of liquid character, consisting of an intimate mixture of two kinds of kinetic units, viz., colloidal particles and smaller units.

5. *Gel*, a colloidal system of solid character, in which the colloidal particles somehow constitute a coherent structure, the latter being interpenetrated by a (usually liquid) system consisting in kinetic units smaller than colloidal particles [1]. If this liquid system is almost entirely or wholly removed the gel will be called a *Xerogel*.

The above list of definitions contains terms which are used in both volumes I and II of this book. It does not yet discriminate between the different possibilities regarding the nature of the kinetic units of large mass. In the next subparagraph these differences will be considered in the case of sols, in more detail. It will be seen that we may discern three cases, according to wich Colloid Science has to deal with the study of three kinds of objects:

> Lyophobic Colloids (dealt with in Vol. I)
> Macromolecular Colloids }
> Association Colloids   } (dealt with in Vol. II).

In these terms the word Colloids stands, strictly speaking, for Colloidal Systems as defined above.

---

[1] Including free colloidal particles possibly present.

In this Volume II the term Colloid will be frequently used not in the sense of a colloidal system, but to designate the organic chemical substance which constitutes the colloidal particle. This is only reasonable in the case of macromolecular and association colloids which form the objects of study in this Vol. II.

Some other terms, also to be used frequently, and which are connected directly with this point of view, are also included in the list of additional definitions which follows:

6. *Colloids*, a name for substances falling within the special sphere of interest of Colloid Science, because
either their single molecules fulfil the definition of colloidal particles (*macromolecular colloids*)
or only the reversible associations of the single molecules (in solution) fulfil this definition (*association colloids*).

7. *Colloid Particles*, kinetic units falling within the range for colloidal particles (definition 2) of substances defined in 6) as Colloids.

8. *Colloid Systems*, a colloidal system as defined in 3), containing colloid particles as defined in 7).

§ 2. COMPARISON OF THE THREE TYPES OF SOLS

a. Hydrophobic sols, the type characteristic of Volume I

In the first volume of this book the so-called hydrophobic sols (or more generally lyophobic sols) have been dealt with.

The general concept of these sols as two phase systems, in which the dispersed phase consists of polymolecular particles stabilised by a surface charge ("capillary electric charge") at the phase boundary, proved useful.

This latter charge prevents effective collisions, i.e., those collisions which give rise to a permanent union. By diminishing the range of the electric repulsion, for instance by adding an indifferent salt, these effective collisions do take place and the sol state is rapidly destroyed.

But, although the capillary electric charge of the original sol particles seems to stabilise the sol state quite perfectly, the stability is in reality only an apparent one. Spontaneously — thus without any added salt — every sol of this type will flocculate, though under favourable conditions this may take several years.

This spontaneous coagulation can be explained, when we consider that in the velocity distribution of the Brownian movement a very small fraction of the particles will have very great velocities. Mutually approaching particles with sufficient kinetic energy will give rise to effective collisions in spite of the mutual electric repulsion. As these unions of particles will very rarely occur, the velocity of flocculation will be so small, that for practical purposes the sol state may make the impression of real stability.

Once a sol of this type has been brought into existence by appropriate means, it has its own line of life. It slowly grows older and older by the above mentioned spontaneous union of extremely rapid particles and finally dies by complete flocculation. In short: *the sol state is principally a non-equilibrium state.*

b. Sols of macromolecular colloids

In this second volume colloidal systems will be treated, which differ fundament-ally from the preceeding ones, in so far as the sol state — and also other states not, or rarely, met with in the case of the lyophobic systems — represent true equilibrium states.

To emphasize the contrast between both groups of colloidal systems let us compare the behaviour of a massive piece of gold and of egg albumin crystals both brought into distilled water. In the former case the piece of gold will not spontaneously divide to form a gold sol[1], the albumin crystals however will go spontaneously into solution.

The kinetic particles of the albumin sol have been shown to be all of the same mass and they do not change their mass as time proceeds[2]. *The sol state is here the true equilibrium state and there is no evidence at hand why the albumin particles should not be described as molecules in true solution in the water.*

Many high molecular substances of biological origin, viz., proteins, polymeric carbohydrates and their natural or industrial derivatives, further the high polymeric chain molecules which organic chemistry has managed to synthesise in the last decades, also form spontaneously true solutions in appropriate solvents.

Organic chemistry has taught us, that not only the last named polymers but quite generally all substances mentioned possess a peculiar structure, characterised by a periodic repetition of the same or closely related atom groups. Molecules of this type have been called *Macromolecules.* Accordingly the above named substances will be named *Macromolecular Colloids*, this term being justified because the weight of the smallest possible kinetic unit in true solution — the macromolecule itself — falls al-ready within the range of particle weights we have arbitrarily chosen as the special sphere of Colloid Science. If needed the prefix macromolecular will be used, so for instance this subparagraph could also be entitled *Macromolecular sols.*

It may happen, that the velocity with which a macromolecular substance reaches its ultimate molecular dissolved state in an appropriate solvent, is much a smaller one than in the above mentioned case of egg-albumin in water. This may be due to the presence of certain resistances, their nature not interesting us here.

Initially the macromolecular substance is then not yet wholly molecularly dis-solved but may contain still larger aggregates. *These incomplete solutions will show changes in properties with time.*

Yet there is this fundamental difference with the first group of colloidal systems (the lyophobic sols) that as time proceeds, the solution tends to reach asymptotically the perfectly dissolved state, which here represents the state of thermodynamical equilibrium.

We may add further that in the perfect state of solution as expressed above the kinetic particles of the macromolecular substance need not all be single molecules. They may be partly associated as well, if only there exists in the bosom of the solution a *reversible equilibrium between single and associated molecules.*

Though the polymolecularity of such molecular associations perhaps might remind one phenomenologically of the particles of hydrophobic sols, from the thermo-

---

[1] As auric acid is accumulated on the surface of the colloidal gold particles, the former substance should be present in our mental experiment.

[2] Apart from hydrolytic cleavage which will slowly destroy the prot-in molecules.

dynamical point of view these kinetic units are in no case to be considered as possessing a phase-boundary. Indeed just as fundamental a thorough knowledge of the properties of phase boundaries is for the treatment of lyophobic sols, just so inadequate would it be here.

The shifting of such equilibria may be at the base of the reduction of the particle weight into halves and other simple fractions which occurs reversibly with certains proteins in solution on changing the pH [1].

c.   Sols of association colloids

A peculiar class of colloid systems is presented by soaps and physico-chemically allied substances. If Na-oleate is put into water it goes spontaneously into solution.

If the final concentration is sufficiently small the soap is practically all subdivided into its smallest possible kinetic units: sodium- and oleate ions.

Sufficiently concentrated solutions however behave as typical colloidal systems. Here a reversible equilibrium between the ultimate kinetic units and large aggregates formed by association exists. The position of this equilibrium is shifted by the mere variation of the final concentration.

Substances which behave in this manner have been called "Concentration variable Association Colloids" or, what is much preferable as also other variables are able to shift the association equilibrium, *Association Colloids*. What characterises them is that the ultimate kinetic particles — molecules, ions — lie quite outside the range of particle weights which we have arbitrarily chosen as belonging to the field of Colloid Science, only the aggregates of them being included.

This state of affairs is not met with in the case of macromolecular colloids, both ultimate kinetic particles and their aggregates falling within this range.

d.   Recapitulation

Colloid Science meets with two main groups of sol states:

1) *The sol state represents principally a non-equilibrium state*, the sols being continuously changing in the direction of flocculation. This case is characteristic of the lyophobic sols of Volume I.

2) *The sol state represents principally an equilibrium state* and can be reached spontaneously starting from the dry colloid and an appropriate solvent. The dissolved substance is present in true solution as single molecules or reversible aggregates of them. This case is characteristic of fairly [2] all sols of Vol. II. We may discern two types:

   a) *Sols of Macromolecular Colloids* in which the smallest possible kinetic units — the macromolecules themselves — fulfil the definition of colloidal particles.

   b) *Sols of Association Colloids* in which the smallest possible kinetic units — the molecules or ions — do not, but only reversible aggregates of these fulfil the definition of colloidal particles.

---

[1] THE SVEDBERG and K. D. PEDERSEN, *The Ultracentrifuge*, Oxford University Press, 1940.
[2] For exception see § 3 d.

§ 3 SURVEY OF COLLOID SYSTEMS IN VOLUME II AND AIDS IN THEIR STUDY

a. Preliminary remark

This § 3, though its main purpose is given in the heading, will at the same time, by its special wording, serve the purpose of stipulating that in this volume Colloid Science assumes the character of physical chemistry in a very wide sense, for the special case of the substances defined above, as *Colloids* (p. 3 § 1 b).

Therefore the special terminology of Colloid Science will not be enlarged here in discussing new kinds of colloid systems. The terms to be used will be simply those of general physical chemistry itself. In consequence even the term Sol is temporarily withdrawn and is replaced by *Solution*, which, as we saw above in § 2, is quite justified in the case of sols of macromolecular and association colloids.

b. Solutions proper

As these colloid systems have already been discussed at some lenght in § 2b and § 2c we can directly turn to the question which auxiliary sciences will be fundamental for the study of macromolecular and associations colloids.

As in their solutions perfect (or nearly perfect) thermodynamic equilibrium exists, it seems at first glance, that general considerations on phase boundary phenomena, so extensively studied and applied in Volume I, are of no use at all for the contents of the whole of Volume II.

Though the generalisation from solutions to all study objects will prove to be inadmissable, at least for the solutions themselves the science of phase boundary phenomena has no importance. It is therefore all the more important that other chapters of general physical chemistry should be consulted, viz., those dealing with solutions, association and, as will be seen below, with other subjects in which equilibria are studied.

Equilibria can be studied in two ways; thermodynamically or molecular-kinetically, in which last case a thorough knowledge of the ultimate kinetic units — the single molecules — may be of primary importance. Especially in the case of macromolecular colloids an intimate knowledge of their structure and configuration opens indeed wide perspectives.

We may conclude that *Organic Chemistry* and *Thermodynamics* are the auxiliary sciences sought for. Therefore chapters on these themes will preceed the systematic treatment of the colloidal systems themselves (Ch. II and III).

c. Two phase systems

A solution of a macromolecular or association colloid is a single phase from the thermodynamical point of view.

Many cases are known in which such a solution may separate into two phases, the one representing a dilute solution of the colloid component, the other being either a liquid or a crystalline solid, both rich in the colloid component.

Under favourable conditions these additional phases may be realized as a homogeneous liquid macro layer or as distinct crystal individuals. The designation "phases" for these liquid or solid colloid systems is justified because each of them represents a

state of perfect (or nearly perfect) equilibrium. Moreover the separation or crystallisation gives rise to an establishment of mutual equilibrium between the two liquid layers, or between the crystals and saturated solution.

Let us once more consider the fundamentals for the study of macromolecular and association colloids.

We state first then that still more chapters of general physical chemistry should be consulted, viz., those on solubility, on phase equilibria, and on crystal structure.

Next we state, that, for the first time in our discussions, a phase boundary makes its appearance, namely that between the two coexisting phases.

Does this imply that special considerations on phase boundaries, as those in Volume I of this book, are here of essential importance?

Thermodynamics indicates, that the influence of surface energy upon phase equilibria may be neglected if the area of the phase boundary is sufficiently small. This indeed being the case in the above mentioned examples, we may answer that for the study of the nature of the single phases or of their mutual equilibrium special phase boundary considerations are of no importance.

In some chapters (IX and XI) nevertheless we are for certain purposes interested in studying the phase-boundary itself. This phase-boundary is, however, as we said already but a secondary detail of no importance at all for the internal state of the two adjacent homogeneous colloidal systems.

d. Apparent single colloid systems

When we follow logically the trend of thought set forth above, we must grant that colloid systems of a composite nature are conceivable, for which it no longer holds that special phase boundary considerations are of no importance.

When for instance in a macromolecular solution, separation of a second liquid or of a crystalline phase sets in, and for certain reasons this new colloid-rich phase remains subdivided in the range of "colloidal dimensions", the total area of the mutual phase boundary will be so large, that surface energy does constitute an intrinsic factor in such composite systems.

In principle the latter do not represent equilibrium systems, though they might at first sight give that impression.

In the course of time they will tend to diminish the total area of the mutual boundary phase (for instance by slow coalescence of the ultramicroscopical drops or by slow recrystallisation), which coarsening of the minutely subdivided phase may have various consequences for the properties of the composite system as a whole (e.g., optical properties).

Here we may only mention one of these consequences. As the surface energy enters also as determining factor in the position of the two phase equilibrium, the composition of both coexisting phases must alter in the course of time. For in proportion as the total area of the phase-boundary decreases, the influence of surface energy on the position of the phase equilibrium becomes less.

Many more consequences could be drawn for such composite colloidal systems, but our first question must be if systems of such a nature do really exist in the field of macromolecular and association colloids. We can answer, that such systems are really encountered (see p. 233-238, Ch. VIII § 1 b - 1 e) So for instance composite systems of two liquid phases, whereby the colloid-rich liquid is ultra or sub-

microscopically subdivided within the liquid poor in colloid. These apparent single systems of composite nature have, as might be foreseen from the presence of a very large internal phase boundary, the character of lyophobic sols. Neutral salts in small concentrations will flocculate these "Sols", the floccules formed existing initially as loosely adhering, very minute drops. By subsequent confluence the degree of subdivision of the colloid-rich liquid phase will diminish until finally a single colloid-rich layer will be formed.

In many cases this confluence is, by the high viscosity of the colloid-rich phase for example, so much retarded that the character of the floccules is not much altered in a reasonable time. If the confluence takes place somewhat faster the "amorphous precipitate" after some time assumes a more granular, or even lumpy and sticky character.

Similar apparent single systems of composite nature in which a crystalline phase is very minutely subdivided in its saturated solution, will also have the character of hydrophobic sols. By adding a neutral salt in small concentration the floccules obtained will represent a "micro-crystalline precipitate". As adhering crystals will in general not tend to coalesce, a diminution of the degree of dispersion of the crystalline phase would here only be possible by recrystallisation.

From § 3 e 1 (see below) it will appear that the division "apparent single colloid systems" does not not possess a sharp boundary between it and the next one called "gels".

e. Gels

These colloidal systems, see definition in § 1 b (p. 2), exist both in macromolecular and association colloids. Here only the most important types of macromolecular gels will be considered. For their realization it seems essential for the macromolecules in question to be of the randomly kinked long chain type [1] (see Ch. XII, p. 483).

e. 1. *Macromolecular gels which are preferable considered as two-phase systems*

Certain gels, as they consist of a cohering mass of highly dispersed flocculation aggregates, can also be classed under the preceding division, representing the solid type of apparent single colloid systems (for fuller information see p. 236, Ch. VIII § 1 d). It will be clear, that for these two-phase gels, special phase boundary considerations may be of use in explaining part of their properties.

e. 2. *Macromolecular gels in which one inclines more to the one-phase concept*

In another type of gels practically no macromolecules are present as free kinetic units, as they all form a coherent network throughout the whole system. One and the same macromolecule, however, takes part in regions of the gel, that can be called crystalline and in other regions that may resemble macromolecular solutions. These "solution regions" form a continuous lacunary system throughout the whole gel, the crystalline regions being dispersed in it, though mutually united by single macro-

---

[1] In many protein solutions another type is met with. Here the macromolecule is somehow relatively tightly coiled, resulting in a definite approximatively globular form (see p. 239, Ch. VIII § 2).

molecular chains. As the order of magnitude of the discontinuities in these gels is very small, it seems presumable that the one-phase concept is here already much more preferable than the two-phase concept. Still there is here a difficulty in characterising the nature of this phase The gel could for instance be considered as a crystalline phase with very extensive lattice disturbances. This conception seems to be of doubtful use, however, for instance in a 2% gelatin gel, where the lattice disturbances would amount to more than 98% of the total volume.

e. 3. Macromolecular gels which are preferably considered as one-phase systems

Here the single macromolecules, though for the greater part of their length freely dispersed in the surrounding liquid, are bound together at certain points by cohesion forces or stronger chemical bonds. Though their character of independent kinetic units is thus lost, the free chain elements of the macromolecules still execute kinetic movements.

The relationship with ordinary macromolecular solutions is close in the case, that the binding between the single macromolecules is caused by cohesion forces between polar groupes, the interpenetrating liquid is of the non-polar type and at the same time constitutes a solvent for the free chain lengths of the macromolecules.

By mentally replacing the original liquid by a more polar one or by an appropriate mixture of solvents, these cohesions bonds may be loosened, without disturbing the dissolved state of the rest of the macromolecules. Then the macromolecules assume the character of free kinetic units, though of course reversible associations between them may be present. Thus the gel liquefies to an ordinary macromolecular solution.

By returning to the original solvent, the mutual binding between the polar groups becomes so firm, that it almost entirely loses its reversible character. We have returned to the gel state. The now resulting gel can thus be considered as the solution of one giant molecule, which latter consists in macromolecules locally bound together. As the limits of this giant molecule and the solvent present coincide, this gel can also be described as an amorphous solid solution.

As in gels of this type, even when stronger chemical bonds exist, no phase boundaries are present, they can be called one-phase gels. It will be clear, that for their study the special phase boundary considerations of Volume I have no importance.

f. Recapitulation

f. 1. *Survey of systems*

In this Volume II of Colloid Science we meet with the following study objects [1] :

a. Solutions proper, constituting a single phase,

b. Two phase systems, the phases being not of very small extension. They fall into the types:

b₁)	liquid / liquid
b₂)	liquid / crystalline solid

c. Two phase systems in which one of the phases is very minutely subdivided. It is argued in Vol. I, Ch. I, that these systems may be treated either as two-

---

[1] Polyphase systems and other kinds of phases (mesomorphic), which are seldom met with, are omitted from this survey.

phase or as one-phase systems but in many cases the two-phase concept will be the more convenient one In a similar sense we distinguish between (d) and (e):

d. Two-phase macromolecular gels, consisting of a cohering mass of highly dispersed flocculation aggregates.

e. One-phase macromolecular gels, representing solid solutions of solvent and macromolecular substance.

f. 2. Aids for their study

In the equilibrium systems *a.*, *b.* and *e.*, phase boundaries being absent (*a.*) or constituting but a secondary detail (*b.*), the specialized chapter of physical chemistry on these boundaries, so fundamental in Volume I of this book, is of no use here. Many other chapters of physical chemistry and their thermodynamical treatment as well as chapters of organic chemistry on structure and configuration of macromolecules are here the foundations on which to base our study.

In the non-equilibrum systems *c.*, and presumably also in *d.*, where the mutual contact area of the phases is very great, phase boundary phenomena do take part in determining their behaviour, and the chapter of physical chemistry rejected for *a.*, *b.* and *e.* now gains an importance comparable with those already mentioned.

The loose connection between the contents of Volume I and II of this book, hitherto based only upon the presence of kinetic units of large mass (See definition "Colloid Science" on p. 2 § 1 b), is here reinforced by the apparent single systems of composite nature *c.* and presumably also by the two phase gels *d.*

§4 ON MENTAL ELIMINATION AS A METHOD OF CLASSIFICATION

a. Ways of treating subjects in Volume II

In developing science predilection as to subjects worked on and as to theoretical considerations favoured, plays a great rôle. This occurs especially in the early stages of development, in more established states of science predilection plays already a minor rôle and in the ideal final stage it would disappear altogether. Science would then constitute a system of knowledge in which only logical order prevails, pre. 'ection being wholly absent. The many logical connections between its various subjects should be represented in a polydimensional system of classification, but as man is obliged to use the wholly inadequate linear order of the written record, a predilection as to the chosen order will still remain in classification.

In colloid science the same applies. In its earlier stages predilections of various kinds prevailed. So for instance the idea that we had to do with a totally new branch of science, further in later stages the idea that all its study objects were disperse two phase systems. Nowadays we must grant that colloid science is not a separate science and that it contains two parts (viz., the contents of Volumes I and II), very loosely bound together by a special chosen wording of a definition (cf. §1). Further that the main points of these two parts must really be incorporated at two different places in the vast body of general physical chemistry.

In connection with the ideas mentioned above it was quite natural that as the guiding principle in experimental work a predilection as to the importance of the "colloid component" in a poly-component system temporarily prevailed. This component would determine the characteristic properties of the total system, and

the other components present would only modify these properties because they "act on the colloidal component".

Much experimental work on subjects belonging to Vol. II has been done and explained, using this predilection for one of the components. Many facts found in this manner have been already translated and treated in terms of general physical chemistry, but many still await such a treatment on a more exact basis. So the latter must still be recorded in Volume II in a rather unsatisfactory way.

In fact we shall meet with chapters in which properties of colloid systems are logically derived from theoretical considerations. In others experimental facts will be the starting point and the aim will be to arrive at ideas how the same chapter could perhaps be rewritten starting from a theoretical point of view.

As in classification of the various subjects we are bound to the inadequate linear order (see above), we must choose a classification principle that has at least a logical foundation. The considerations of the next two subparagraphs will be of help in such a choice. Therefore the question of classification is postponed to § 5.

b.   The elimination procedure, case of sols

As we said above, in exact science there is no place for *a priori* predilection for one of the components of a polycomponent system, they all being quite as important or unimportant initially.

All components of a colloidal system should be taken into consideration and only after a complete and quantitative theoretical treatment could it be decided which components under special conditions have little influence and can be neglected, and which others have therefore a great importance.

In many cases the result will indeed be that the colloid component is of great importance. So the *a priori* predilection had truly a heuristic value.

We shall now discuss the point that even a predilection for the colloid component logically driven to the extreme — a backward step from the exact point of view — may have heuristic value in this Volume.

This procedure consists in considering only the kinetic units of the colloid component, all other components present in the given colloid system being mentally eliminated. This procedure will be called *"Elimination Procedure"*.

Now let us take as example an egg albumin hydrosol. From the general physical chemical point of view this colloid system is simply a *solution.*

If we mentally eliminate the kinetic units of all non-colloid components present — here only the water molecules — there remains a system of moving and colliding kinetic units — the egg albumin molecules —, that in many points has the properties of a *gas.*

This result is of course not startling at all, because physical chemistry has already shown us long ago that also in solutions which contain only small molecules, for instance sugar in water, the dissolved substance behaves in the solvent as if it were present as a gas (VAN 'T HOFF, laws of dilute solutions).

Colloid science has also long ago recognised the validity of this principle for the kinetic units in which it has a special interest and has based its various osmotic methods of determining the molecular weight of colloids on it, using the gas laws.

Indeed the principle holds not only for small dissolved molecules but also for much larger kinetic units and even for microscopically visible particles suspended in a liquid (PERRIN).

The elimination procedure, which in the case of sols did not give new points of view, however does so when we apply it to polyphase systems which occur in the domain of macromolecular colloids.

c. The elimination procedure applied to various colloid systems of macromolecular colloids

In the following discussion we shall use the term *macro-units* for the kinetic units fulfilling the definition of colloid particles (p. 3 § 1 b) and *micro-units* for those smaller than them. Almost all colloid systems of macromolecular colloids consist of an intimate mixture of macro- and micro-units, and we shall now apply the elimination procedure to them.

c. 1. *Solutions proper*

As already shown above, after mental elimination of the micro-units present, there remains a system of macro-units — here the macromolecules and (or) their reversible aggregates — with the character of a gas. The system obtained may be called a *colloid gas*.

c. 2. *Two phase systems of the type liquid/liquid*

Under various conditions, macromolecular solutions may separate into two liquid layers, the one poor in colloid, the other rich in it. In colloid science this partial miscibility has been called *Coacervation*, and the colloid-rich phase *Coacervate*.

As a coacervate fulfils the definition of "Sol", given in § 1 b (p. 2), we could also simply say that both layers are *"Sols"*, thus constituting an equilibrium between two sols of different colloid concentrations.

But then we should come into conflict with the above result, that a sol can be considered as a colloid gas. For as gases are miscible in all proportions, why then would it not be the case with two colloid gases of the same substance, but of different concentrations?

Thus we already get the impression that the term Sol seems inappropriate for both layers.

Indeed, if we mentally eliminate all micro-units in the two-phase system, we shall see two spaces, the one — corresponding to the colloid poor layer — with relatively few freely moving macro-units, the other, corresponding with the colloid-rich layer, filled with many macro-units, closely packed though still in movement.

The system in the first named space can certainly be considered as a dilute *gas*, that in the second space can be either a very compressed gas or liquid. Taking now into consideration that a boundary between both systems of macro-units is permanently present and the total system represents an equilibrium state, then we conclude that the state of the macromolecular substance in the second space may be compared to a *liquid*.

The latter system, corresponding to the coacervate layer, may be called *a colloid liquid*. As to the nature of the equilibrium between both systems of macro-units, we can characterise it as a *liquid-vapour equilibrium of the macromolecular substance*.

We thus see that the elimination procedure firstly gives a certain justification for introducing a separate term — *coacervate* — for the colloid-rich layer, secondly opens perspectives for studying macromolecular two phase colloid systems of the type liquid/liquid from general physical chemical points of view.

c. 3. *Two phase systems of the type liquid/crystalline solid*

In the same way a two phase system of the type liquid/crystalline solid offers, after mental elimination of the micro-units present, the picture of a *sublimation equilibrium between saturated vapour and crystals of the macromolecular substance.* These crystals which differ in so far from the original crystals, that micro-units possibly present in it are mentally eliminated, may be called *colloid-crystals.*

c. 4. *One phase and two phase gels*

In discussing the structure of these gels in § 3 e., the elimination procedure was already used unconsciously in some form. We need therefore not repeat here the points of view obtained.

d. Recapitulation

1. Predilection *a priori* for the macromolecular component in a poly-component sytem, common in earlier stages of colloid science, should in the exact treatment of ideal colloid science no longer occur. Nevertheless its heuristic value might in a given case receive its theoretical justification *a posteriori.*
2. Predilection in the order chosen for the written record, seems unavoidable also in ideal science.
3. Predilection for the macromolecular component driven to the extreme by eliminating mentally all other components may have a heuristic value:
   a. sols then appear as macromolecular gases;
   b. coacervates as macromolecular liquids;
   c. the two phase equilibrium of the type liquid/liquid as evaporation equilibrium of the macromolecular substance and
   d. the two phase equilibrium of the type liquid-crystalline solid as sublimation equilibrium of the macromolecular substance.
   The *a posteriori* theoretical justification exists as yet only in the case of a.
4. In the course of this chapter three different liquid macromolecular colloid systems have been met with, which fulfil the definition of Sol in § 1 b (p. 2):
   a. sols proper; b. coacervates; c. a group of apparent single colloid systems (p. 7 § 3 d).

§ 5. CLASSIFICATION OF SUBJECTS

a. Choice of a main classification principle for colloid systems

We wil now continue the discussion on classification interrupted in § 4 a. As for all subjects in Vol. II equilibria are of great importance (cf. § 2 and 3), a general thermo-dynamical treatment, including application of the phase rule seems obvious.

Nevertheless as classification principles those which are indicated for the phase rule will not be adopted here for a classification of subjects. These principles are, as

is well known, the number of components and the number of phases, for these two only occur in the phase rule itself.

In textbooks on the phase rule the first one — the number of components — is chosen as the main classification principle. As to the kinds of phases the phase rule has, strictly speaking, hardly any interest at all. The phases are only labelled as gaseous, liquid, nematic, smectic, or crystalline, to indicate that they are *different*, so that they can be counted and thus the number of phases can be obtained.

In this Volume of Colloid Science, we shall in general have no interest in numbering the different colloid systems. Indeed we are in the first place specially interested in each of the kinds of phases that can occur and especially in their internal structure.

For we want to bring the properties of each of these phases into direct relation with the structure of the kinetic units present in them.

Thus we will seek to explain the properties of a colloid system from those of the macro-units and micro-units present and from their interaction. In second place comes the interest for two and more phase equilibria and though the phase rule will help in treating them in a general but formal way, the aim of this book will nevertheless once more be to explain these equilibria from considerations regarding the internal states of these phases, starting from the macro and micro-units present and from their interrelations.

But all these points are only superfluous details for the phase rule.

It wil thus be clear that as the main classification principle one of those [1] used in the application of the Phase Rule will not be appropriate, but on the contrary the classification must be based on the different *kinds of phases*. Having once adopted this it follows that for further subdivisions the number of phases may next be used and that the number of components present will be of least importance.

b. Macromolecular colloids

For practical and theoretical reasons explained later in § 5 c (p. 17) the macromolecular and association colloids will be dealt with separately. This facilitates our task of finding a satisfactory classification of subjects for the former group. As the smallest possible kinetic unit, the macromolecule, already falls within the sphere of colloid science, the classification can be based on the macromolecule, and the systems formed by it. Thereby we make use of the elimination procedure (p. 11 § 4 b) and take into consideration the conclusions at the end of the foregoing subsection.

---

[1] Apart from the essential objection made others are also present. For instance in choosing the number of all components present as the main classification principle, the classification system will be utterly impractical and in many cases not applicable at all. Indeed it will be impractical to seek in three different main divisions of the book if we will get an idea how different salts will modify certain properties of, for instance, Na-arabinate dissolved in water. Without added salt we m"st seek in the main division treating two component systems, with $NaCl$ added in the main division treating three component systems and with $CaCl_2$ added once more in another main division treating four component systems. Of course this impractical aspect can be removed if we apply the elimination procedure. Then all three solutions are "one-colloid component systems". But even basing the classification system only on the number of colloid components present, it will still give serious difficulties. For in the field of macromolecular colloids very often an apparently single colloid consists of a large and unknown number of separate colloid components, which differ in the chain length of the macromolecules. Thus it would not even be possible to count the number of colloid components.

We shall thus, in a first division, consider the macromolecule itself, its mass, constitution, configuration and physical properties. In this main division nevertheless solutions of macromolecules and some of their properties already enter into the discussion based upon thermodynamical and rheological considerations on a model of the randomly kinked macromolecule.

In the next main division the different kinds of single colloid systems will be treated. Here once more solutions are dealt with as a subdivision, but now the stress is laid on the actual properties of the whole system, including interactions between macromolecular ions among themselves or macromolecular ions and micro ions.

For the classification of the different kinds of single colloid systems the elimination procedure is used to characterize them. It enables us to divide the liquid single systems into two kinds, which will be called *Sols* (in the restricted sense) and *Coacervates*, though both fulfil the definition of sol in the larger sense of § 1 b (p. 2).

As the number of phases is of smaller importance to us than the kinds of phases (p. 14 § 5 a.), the two phase equilibria coacervate/sol will be found in the Chapters VIII and X dealing with coacervates, likewise the two phase equilibria crystalline state/sol in the Chapter VIII dealing with the crystalline state.

In the third main division will be found the apparent single colloid systems and in the fourth gels, containing an interpenetrating liquid system of micro-units. The last main division will deal with solid systems of a nature not further specified, consisting only of macromolecules.

For further details and reference to the corresponding chapters, see the section entitled Part I of the survey on page 16.

The actual contents of the consecutive chapters do not slavishly follow the classification given in this survey.

CHAPTER VIII, entitled: "CRYSTALLISATION, COACERVATION, AND FLOCCULATION brings together three different divisions of the scheme (2 B *a*, 2 C and 3).

CHAPTER IX, entitled: REVERSAL OF CHARGE PHENOMENA, EQUIVALENT WEIGHT, AND SPECIFIC PROPERTIES OF THE IONISED GROUPS, not only concerns the phase boundary of coacervates, but also the boundary of floccules and even the boundary of adsorbed colloid films on particles (e.g., on SiO_2).

CHAPTER X, entitled: COMPLEX COLLOID SYSTEMS, though for a greater part dealing with coacervates, also takes into consideration colloid systems of other kinds in which the ionised groups play a similar rôle.

CHAPTER XI, entitled: MORPHOLOGY OF COACERVATES, consideres in addition colloid bodies which, although they no longer consist of coacervates, have been produced from them.

Still more grave objections could be made from the point of view of rigid classification, as in the last named three Chapters IX, X and XI, colloids have often been considered (phosphatides) which do not at all belong to macromolecular colloids, but to association colloids, the latter belonging to Part II of this Volume.

c. Association colloids

If we were free to write a general physical chemistry of soaps or soaplike substances — thus not hindered by the necessity in this book of paying only special attention to particles falling within an arbitrarily chosen range of masses — then an analogous classification as above for macromolecular colloids would suit us.

Scheme for the classification of Subjects in Volume II

PART I. MACROMOLECULAR COLLOIDS

1. *The macromolecule itself* (also in dissolved state and in solid matter)
    A. formation and structure   Ch. II
    B. thermodynamics   Ch. III
    C. physical properties   Ch. IV
    D. molecular weight   Ch. V

2. *Single systems of macromolecules + micro-units*
    subdivided as to the nature of the system of macromolecules that remains after mentally eliminating the micro-units, in gaseous, liquid and crystalline macromolecular systems.
    A. "gaseous" macromolecular systems: SOLS IN THE RESTRICTED SENSE, for practical purposes divided into two separate chapters:
        the macromolecule is a non-electrolyte   Ch. VI
        the macromolecule is an electrolyte   Ch. VII
    B. "liquid" macromolecular systems: COACERVATES
        α) coacervates, in which the ionised groups of the macromolecule, if present, play no rôle   Ch. VIII
        β) electrical properties of the coacervate boundary giving information on equivalent weight of the macromolecule and specific properties of the ionised groups   Ch. IX
        γ) coacervates in which the ionised groups do play a significant rôle
                                                    Ch. X
        δ) coacervate bodies (i.e., small amounts of coacervate limited by their natural coacervate boundary, see p. 17 § 5 d).
                                                    Ch. XI
        These objects unite the elements, which, idealized as of infinite extension, are studied separately in Chapters VIII and X (the three dimensional interior), and in Chapter IX (the two dimensional boundary).
    C. "crystalline" macromolecular systems: COLLOID CRYSTALS
                                                    Ch. VIII

3. *Apparent single systems of macromolecules*
    being higher dispersed systems of two (or more) kinds of the single systems of division 2.
    Some cases (the total system has a liquid character, or consists of a flocculated mass) are described in   Ch. VIII
    This division does not possess a sharp boundary between it and the next one.

4. *Gels, containing an interpenetrating (usually liquid) system of micro-units*
    which comprise systems which may preferably be described as *two phase gels* and others which are better described as *one phase gels*.
                                                    Ch. XII

5. *Solid systems, consisting only of macromolecules*
                                                    Ch. XIII

PART II. ASSOCIATION COLLOIDS

The number of colloid systems considered is here restricted to only two: *sols and coacervates*   Ch. XIV

We should only have to change the term macromolecule into, for instance, soap molecule and the term colloid system into soap system.

As however we are restricted, we must select from that classification those divisions in which large soap-aggregates occur as kinetic units. Let us review the five main divisions in Part I of the survey on page 16.

Div. 1. The single soap molecule falls already outside the direct sphere of colloid science.

Div. 2. It seems probable that in the crystalline phases only single soap molecules and not their aggregates constitute the architectural units. It is difficult to decide whether in the nematic and smectic soap systems the aggregates form the constitutional units or not. Because of this lack of experimental information only *sols* and *coacervates* remain in this division.

Div. 3 and 4. Certainly soap systems belonging to these divisions are known. For instance coarse structured gels in which a network of fibres can be distinguished ultramicroscopically and on the other hand clear soap gels which are probably "one phase gels". As however our knowledge of these systems is limited we shall not deal with them in detail.

Div. 5. Solid systems consisting only of soap molecules lack, of course, the characteristic units of association colloids, the ultimate structural units being only soap and alkali ions.

So there remain only a very restricted number of subjects analogous to those in macromolecular colloids. Moreover we have already discussed the fact that with the association colloids only the associated low molecular units form the kinetic units of the colloidal system.

Thus association — and not the molecule — forms the basic idea of this class of colloids. Therefore in the chapter on association colloids (Chapter XIV) no effort is made to follow an analogous division into subjects as in macromolecular colloids. Indeed the association phenomenon itself is put in the centre of the discussion and an attempt will be made to explain the properties of sols and coacervates from this point of view.

d. Static and dynamic colloid morphology

Up till now we directed our attention especially towards the three dimensional interior of the different single kinds of phases, without considering the actual extension of the phases in space.

In reality a phase is always limited, a sol for instance artificially by the walls of the containing vessel and by the air.

But in the case of coacervates and crystalline phases suspended in their equilibrium sols, as small drops or crystal individuals, we meet with study objects having a natural boundary. These objects will be called *Colloid Bodies*. They may have characteristic shapes, resulting from the fact that the three-dimensional phase is present in a limited amount. A colloid body thus can attract our interest because it is a three-dimensional phase, limited by a two-dimensional boundary.

The aim of "*Static Colloid Morphology*" is to explain the actual shape of a limited amount of a colloid rich phase from the properties of both elements. The case considered is of course the simplest one. More complicated study objects are for instance colloid bodies of a composite nature, consisting of more than one colloid rich phase.

A far more extended field is that of *"Dynamic Colloid Morphology"* which registers and purports to explain the morphological changes of a given colloid body, when for instance one or more of the variables determining the original equilibrium (e.g., temperature, concentration or kinds of micro-units present) are changed.

Here a new factor enters, *the actual path that the system follows towards a new equilibrium state,* this path being the cause of the morphological changes. From a knowledge of the equilibrium states alone, the path which will be followed as a function of the time cannot as yet always be predicted.

Colloid morphology, especially its dynamic variant, is still in its very beginnings, but the author thought that a survey of results obtained with coacervates (Ch. XI) might stimulate its further development.

II. THE FORMATION AND STRUCTURE OF MACROMOLECULES

R. HOUWINK

Rubberfoundation, Delft.

---

§ 1. THE EXISTENCE OF MACROMOLECULES AND THEIR BEARING ON COLLOID SCIENCE

The origin of the conception of macromolecules can be traced back[1] to 1863, perhaps even as far as 1846 (BAUDRIMONT, see Vol. I), but only in later years has it become a lively subject of study. Mentioning among others the names of EMIL FISCHER, FREUDENBERG, HAWORTH, MARK, K. H. MEYER, SPONSLER, and DORE, the greatest honour however, must be bestowed on STAUDINGER[2], who founded the conception of macromolecules in a long series of publications.

By this expression, which itself contains a certain contradiction in terms, very large molecules of from 1250 Å upwards are implied, consisting of a pattern of regularly repeated structural units, i.e., the so-called monomeric groups or ground molecules. The paraffin chain in Fig. 1 represents one of the simplest types, with CH_3- as terminal groups.

$$CH_3 \ldots\ldots CH_2 - (CH_2)n \ldots\ldots CH_2 - CH_3$$

terminal   ground
group   molecule
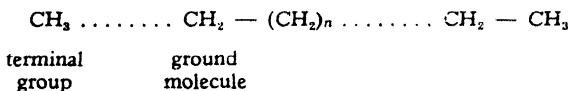
Fig. 1. Paraffin chain, one of the simplest macromolecules.

Considering chain structures only, the macromolecular conception is sound, but difficulties arise when also taking globular molecules into consideration. Thus, in Fig. 2 where the molecule must be thought of as extended in all directions including above and below the plane of the paper, the whole conglomerate can be considered as one large macromolecule. It is usual in macromolecular chemistry to argue that all groups attached to each other by primary bonds (see p. 21) should belong to the molecule, the latter being attached to its neighbours merely by the much weaker secondary bonds.

---

[1] LOURENÇO, *Ann. Chim. et Physique*, 67 (1863) 275.
[2] For a complete summary of his older works, see H. STAUDINGER, *Die hochmol. organ. Verbindungen*, Berlin 1932.

$$
\begin{array}{c}
\vdots \qquad\qquad \vdots \\
| \qquad\qquad | \\
CH_3{-}C{-}CH_2{-}CH_2{-}CH{-}CH_2{-}\ldots \\
| \\
CH_2 \\
| \\
\ldots{-}C{-}CH_2{-}CH{-}CH_2{-}CH_2{-}\ldots \\
| \qquad\quad | \\
CH_2 \qquad CH_2 \\
| \qquad\quad | \\
\ldots{-}C{-}CH_2{-}C{-\!-}CH_2{-}CH{-}\ldots \\
| \qquad\quad | \qquad\qquad | \\
\vdots \qquad\quad \vdots \qquad\qquad \vdots
\end{array}
$$
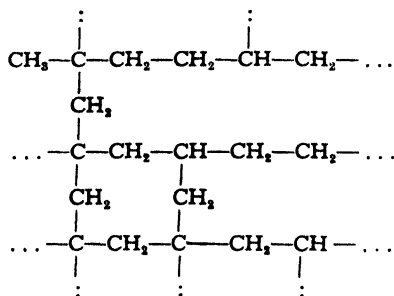
Fig. 2. Three-dimensional macromolecule.

However, to remain consistent, substances like diamond and Fe_2O_3 crystals must also be called macromolecules, a line of reasoning often followed in macro-molecular chemistry. This, however, brings us into serious conflict with classical chemistry, in which just one Fe_2O_3 group, cut out from a crystal, is considered as *the* molecule.

There is no easy way out of these difficulties, both lines of reasoning being justified to a certain extent. For practical reasons we will consider a substance as macromolecular, when, *in its final form or during some stage of its formation, separate kinetic units of macromolecular dimensions are to be distinguished.* Thus, a piece of hardened phenol-formaldehyde or of undercooled silicate glass will both be considered as macromolecular. The former because during polymerisation, and the latter because during cooling (in the transformation interval, see p. 654), large separate particles have existed, cohering by means of primary bonds. If this procedure was not followed, one would come to the inconsistency that in the state of transition the substance would be called macromolecular, whereas on increasing the size of the molecules still further (by polymerisation or cooling) the macromolecular state would be abandoned again. Fe_2O_3 however, crystallises suddenly on cooling its melt, there being no state of transition in which separate particles of macromolecular dimensions are present. On this practical ground it is useless to talk of macromolecules being present, even in the final crystallised state.

We are aware that this procedure is rather arbitrary and open to objections; nevertheless it has been adopted as a working basis.

It is without doubt that, since the introduction of the macromolecular conception, colloid science has assumed quite new aspects. Many properties of colloidal solutions and of solid colloids can hardly be explained nowadays without the assumption of very large molecules. Among these may be mentioned the low osmotic pressure, the fact that the melting point and boiling point are no longer characteristic, and the great delay in reaching equilibrium conditions in various processes (slow dis-solution) as a consequence of the retarded diffusion. For a survey of these phenomena Chapter I in Vol. I, by KRUYT, may be referred to. Many of the older theories, e.g., those explaining the high viscosity of rubber and cellulose solutions with the aid of very thick mantles of adsorbed solvent, have been abandoned and fresh theories advanced in order to explain plastic and elastic properties of all kinds of natural and synthetic products.

§ 2. BONDS IN AND BETWEEN THE MACROMOLECULES[1]

In the foregoing section the definition of the macromolecule was connected with the existence of primary and secondary bonds and therefore it is necessary to give a more precise description of these.

Both kinds are distinguished from each other by their energy content. For the primary bonds this is of the order of 100 kcal/mol or more; for the secondary bonds, of the order of 10 kcal/mol, although no sharp distinction can be made, there being transitional cases.

The *primary bonds* include two types, the homopolar and the heteropolar. The first can be found in C—O or in C—H and is based on the sharing of two electrons by the attracting atoms. The heteropolar bond is found, for instance, in NaCl and is caused by the electrostatic attraction between the Na· and the Cl′ ions. The energy content of both types is so high that they cannot be loosened easily by heat or by mechanical force, and thus a substance, cohering in all directions by means of primary bonds, cannot be plastified and has a high weakening point and a great mechanical strength.

For the *secondary bonds* the situation is much more complicated, and one can distinguish three main types. For each type, the mutual attractive forces r decrease, in a different way, as the distance between the attractive centres increases, according to the formulae mentioned below, A, A', etc being constants.

a. Attraction by permanent dipoles
$$\tau = \frac{A}{r^4}$$

b. Attraction as a consequence of polarisation of a neutral atom or atomic group (induction effect) under the influence of a dipole
$$\tau = \frac{A'}{r^7}$$

c. Attraction as a consequence of the dispersion energy as such (Van der WAALS' LONDON attraction)
$$\tau = \frac{A''}{r^7}$$

If the decrease is slow, the forces have a far-reaching effect and the molecules can be separated relatively far from each other without loosing their coherence. Such substances as a rule will be less rigid than those with short-reaching forces.

There is a principal difference between the dipole forces (including the induced dipoles) and the dispersion forces, in so far that the former have a directional effect, while the latter is non-directional. Consequently the dipole forces will lead more easily to crystallisation phenomena.

Here one may consider the *mesomeric state*[2], a condition of maximum stability for compounds, containing π electrons as a double bond. This double bond may be conjugated with another double bond or connected with an atom having an available lone pair of electrons (O, N, S, halogens). Then one or more of the "unperturbed structures" of the molecule (the individual structures, which are capable of representation) show dipole forces, the strength of which is intermediate between normal dipole

[1] A survey is given by H. MARK, *Physical Chemistry of High Polymeric Systems*, New York 1940.
[2] For a survey see H. B. WATSON, *Modern Theories of Organic Chemistry*, Oxford 1937.

forces and heteropolar forces. These unperturbed structures may be encountered in compounds containing $-C\underset{NH}{\overset{O}{<}}$ groups and to a lesser extent in aromatic compounds and in compounds containing $>C = O$ and even $>C = C<$ groups. These dipole forces play an important part in proteins and, in diminishing order, in high-molecular aromatic compounds, ketones, esters and perhaps in macromolecular unsaturated compounds.

A particular type of mesomeric phenomenon (also, sometimes explained by means of electrostatic attraction) is the so-called *hydrogen bond*. It appears probably that the bonding is due to a hydrogen atom resonating between two atoms — most often between two oxygen atoms or an oxygen atom and a nitrogen atom.

Such a hydrogen bond can, for example, link together two OH groups in cellulose, and play a part in proteins, where neighbouring —CO—NH— groups can be linked together by hydrogen bonds between the oxygen atom of one group, and the nitrogen atom of another group.

Apart from these attractive forces, there are always *repulsive forces* acting between the atoms, due to the mutual repulsion of their electrons and of their nuclei. They have a very short range of action so that they are only observed when the atoms approach each other very closely; this is the resistance we feel, when compressing a solid.

The difference between the attractive and the repulsive energies leads to a resultant, expressed in the potential curves, discussed by KRUYT in Vol. I.

§ 3. THE FORM OF THE MACROMOLECULES

Macromolecules can be roughly divided into three groups according to their form.

a. *The linear or one-dimensional polymers*, where the molecules have grown into a *chain- or thread-form*. It is not necessary that these chains are stretched; on the contrary there is much more evidence that they are usually coiled up clews. The density of these coils depends on the mutual attraction between the links of the chain and on the ease with which these links can rotate with regard to each other. This coil-form is so important from a colloidal point of view that in the next chapter it will be treated at length.

b. The second group contains the globular *or three-dimensional* molecules (Fig. 3 c), which have become enlarged in all directions during polymerisation. There are certain rules[1] with the aid of which one can predict whether polymerisation will lead to linear or globular molecules. If the growing molecule has two active spots only, i. e., if it can only grow at two places on its surface, chains will be formed. When there are, however, three or more active spots on the surface, then globules are to be expected.

c. A third group is formed by *nets*, which to a certain extent can be considered as an intermediate between chains and globules, as appears from Fig. 3 b. Here, this formation starts from existing chains, e.g., those supplied by nature, which are later interlinked by what may be called chemical bridges. In the case of Fig. 3c sulphur

---

[1] R. H. KIENLE, *Ind. Eng. Chem.*, 22 (1930) 590.
[2] It will appear on p. 172 that the most probable form of the coils is that of a bent bean.
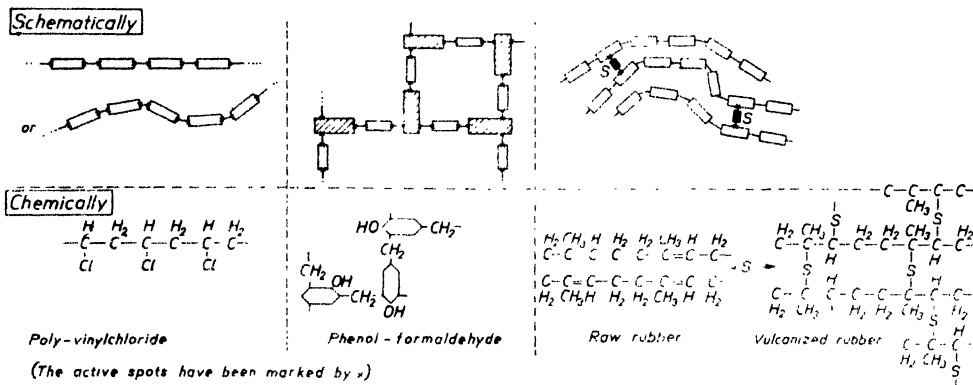
a Chain formation. *b* Formation of three-dimensional structures (globules) *c* Net formation (interlinking between existing chain molecules.

Fig. 3. The three chief principles for building-up macromolecules.

forms a bridge during vulcanisation of rubber. In practice such nets will be found to be folded together.

At first sight there seems to be no large difference between the three groups, because the linear polymers form a more or less [2] globular coil, the three-dimensional polymers are globular, and the nets are folded together, which ultimately must also lead to a more or less globular form. The difference however, is that in the linear polymer the chains are not interlinked, so that the coils can be spread out widely [1] and in extreme cases they can even be stretched into parallel bundles. The globular polymers are so intensively interlinked that the globules hold together strongly, and by no physical means (except breaking the molecules) can an important change of form be effected. In swelling liquids it may be that some extension to 2 or 3 times the original volume takes place, but there is no question of a complete unwinding as with chain molecules. For nets the situation is intermediate as already stated. There is a certain interlinking, *but this is to a lesser extent than with globular polymers*. The net can therefore be easily deformed and swelling media will cause an increase of volume to many times the original volume. When a net becomes very intensively interlinked (which is sometimes possible, e.g., when ebonite is formed), the properties of the material will approach those of a globular polymer.

These brief considerations are already sufficient to stress the importance of the molecular form in colloid chemistry.

§ 4. THE DIMENSIONS OF MACROMOLECULES

It is rather arbitrary whether a molecule shall be considered as a macromolecule. Following the system by STAUDINGER [2], based on the degree of polymerisation, one arrives at the classification of Table 1.

---

[1] M. BERGMANN and C. NIEMANN, *J. Biol. Chem.*, 118 (1937), 301; *Ibid, Science*, 86 (1937) 187.
[2] A recent summary: H. STAUDINGER, *Organische Kolloidchemie*, Braunschweig 1940.

TABLE 1

CLASSIFICATION OF MOLECULES ACCORDING TO STAUDINGER

	Degree of polyme- risation	Number of atom groups in the chain	Length of the shortest chain-mol. of this type (paraffin) in A	Lowest mol. weight of this type (paraffin)	Examples
Low-molecular	< 10	< 50	< 25	< 128	Canada balsam, colo- phony, resin.
Hemicolloidal	10—100	50—100	25—250	142—1402	Polyfructosane, wood- polyose, hemicolloidal polystyrene.
Mesocolloidal	100—500	500—2000	250—1250	1402—7002	Lichenin, cellulose rayon, intensely plas- ticised rubber.
Eucolloidal (Macromolecular)	> 500	> 2000	> 1250	> 7002	Cellulose, rubber, po- lystyrene.

According to this a chain molecule is called macromolecular when its length is greater than 1250 Å.

The question is whether STAUDINGER is right in his nomenclature of hemi-, meso-, and eucolloids. On the basis of the classification by OSTWALD [1] where the limits of colloidal dimensions are fixed between 10 and 10^4 Å, a solution containing hemi- or mesocolloids would in fact be colloidal. Amongst the eucolloids, however, one often meets substances in which the chain molecules have a length of 10^4 Å or more, and in such cases OSTWALD's upper limits would be surpassed, whereas STAUDINGER still considers the molecules as colloidal. There is thus a certain dis- crepancy between the two nomenclatures.

Against STAUDINGER's system of classification as a whole it may be objected that it has been created for chain molecules only, neglecting globular particles. Thus a protein with a molecular weight of about 10^7 e.g., (haemocyanin), if it had extended chain molecules would reach a length of over 500 000 Å, whereas its diameter would be only of the order of 300 Å if it were globular [2]. The chain-form therefore would bring this molecule far above the upper limit of colloidal dimensions (10^4 Å), whereas the globular form would just make it a colloid.

In our opinion the best way to decide whether a particle shall be considered as colloidal is to keep to the definition given by OSTWALD, with this alteration however, that a *particle will be considered as colloidal if at least one of its dimensions is between 10 and 10^4 Å, none of its dimensions being larger than 10^4 Å*. We agree that it is rather dogmatic to keep to such limits and therefore this system will only be used as a very general guide.

In this classification rubber-latex, having particles of about 1 μ ($= 10^4$ Å) is still a colloid, but reaches to the upper limit. This is in agreement with the fact that the latex globules can be observed microscopically only just to exhibit Brownian movement.

---

[1] See Vol. I, Chap. I (KRUYT).
[2] This can be calculated, starting from the specific volume (0.75 for proteins). A volume of 24 000 Å is obtained for particles with mol.wt. = 17 000.

TABLE 2

SIZES OF SOME IMPORTANT KINDS OF MACROMOLECULES

	Rubbers	Polysaccharides and connected materials	Proteins	Resins	
				linear	globular
Hemicolloidal (pol. degree 10—100)	Rubber, broken down to a syrup (Rubbone)	β-cellulose	—	Polystyrene, polyvinyl chloride etc. in the liquid state	Colophony, phenol- or urea-formaldehyde in the solid A-stage
Mesocolloidal (pol. degree 100—500)	Normally plasticised rubber, native balata, and gutta percha	α-cellulose, rayon, cellophane-foil	Pepsin Insulin	ditto, in the solid state; sometimes rubbery	Shellac (probably), somewhat poly-merised damar and copal. Phenol- or urea-formaldehyde between the A- and B-stages.
Eucolloidal (pol. degree > 500)	Natural rubber, synthetic rubbers (neoprene, Buna, Butyl-rubber, Thiokol)	Natural cellulose fibres (wood, cotton, flax, hemp); starch; pectin.	Keratin CO-hæmoglobin Serum albumin Hæmocyanin	ditto, in the highest polymeri-sation stage	Polymerised shellac, damar and copal. Phenol- or urea-formaldehyde between the B- and C-stages

Starting from the definition just given, the fact is encountered that not only linear polymers, but globular also[1], can cover *a far greater field than that according to the colloidal dimensions*. As this book is devoted to colloid science, we will confine ourselves therefore chiefly to macromolecules up to 10^4 Å; only in special cases some reference will be made to larger ones.

In order to give some indication of the size of important macromolecules, the figures[2] of Table 2 may be mentioned. Here the STAUDINGER classification into hemi-, meso-, and eucolloids is still used, although this loses most of its merits when based upon dimensions in Å instead of using the polymerisation degree as a criterion.

It appears from this table, that for most products mentioned, representatives can be found in the low-molecular and the high-molecular state. Here again we meet with the great number of possibilities from a colloidal point of view.

§ 5. CHEMICAL CONSTITUTION

From the point of view of production, four groups of macromolecules can be distinguished, namely:

A   Those formed by nature;
B   Those derived from natural macromolecules by modification (derivatives);
C   Those entirely built up along synthetic lines;
D   Inorganic macromolecules.

We will consider the first two groups together, because they are so closely connected with each other; in group C the so-called *co-polymers* will be encountered, in which two ground molecules x and y are alternatively attached to each other as follows:

$$\ldots\ldots x - y - x - y - x - y \ldots\ldots$$

For an example of this one may cite GRS in Table 8 on p. 37. The group of inorganic macromolecules will be considered later.

A and B   Natural macromolecules and their derivatives
1.   Polysaccharides[3].

Without doubt the *polysaccharides* form one of the most important groups of macromolecules, because *cellulose* and *starch* belong to this group. It is a striking fact that the chemical formulae of the various polysaccharides are so much alike, as may be seen from Fig. 4. Cellulose and starch, for example, appear to be only stereo-isomers of the $C_6H_{10}O_5$-group. On the other hand, considering their large difference in properties, it will be clear that there must be special causes for this (see p. 27). Other important members of the polysaccharide group are glycogen, inulin, mannane, galactane, and pentosane.

---

[1] On p. 672, for example, it will be shown that a macroscopic piece of certain resins can be polymerised into one large macromolecule.

[2] When not otherwise indicated, the molecular weights have been measured by means of the viscosimetric method. In Chapter VI, it will appear that it is still doubtful whether this method measures the exact magnitude. As a means of comparison however, it can be used.

[3] W. N. HAWORTH: various publications in *J. Chem. Soc., London*, since about 1925. Summaries: E. F. and K. F. ARMSTRONG, *The Carbohydrates*, London 1934; K. H. MEYER, *Natural and Synthetic High Polymers*, New York 1942 II; H. PRINGSHEIM, *Polysaccharide*, 3. Aufl., Berlin 1931.

β-Glucose

α-Glucose

Cellulose

Amylum (Starch)

α-Mannane

β-Galacturonic acid
(R = CH₃ or H)
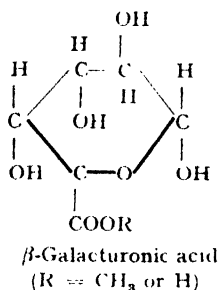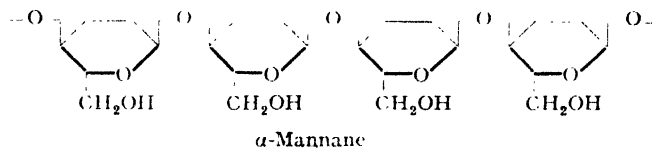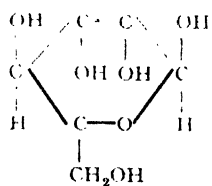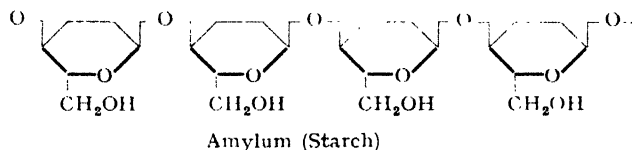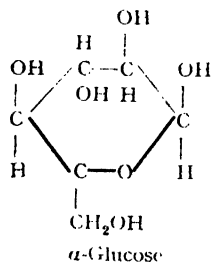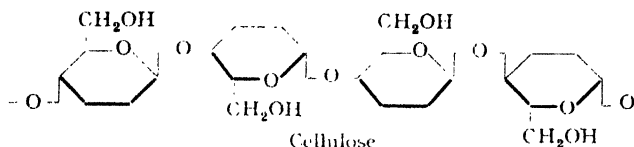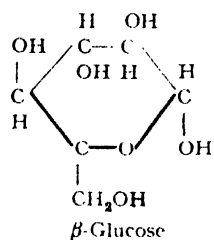
Polygalacturonic acid
This acid or a combination of it with other polysaccharides is an
element of pectin.

Fig. 4. Formulae of some polysaccharides.

Closely connected with these pure polysaccharides is *chitin*, built up from glucosamine (glucose with one NH_2 group). *Pectins* seem to be either polygalacturonic acid itself, or a combination of this acid with polysaccharides, like arabane or galactane. This polygalacturonic acid contains a COOH group instead of CH_2OH, and thus one can expect the formation of salts.

Among the *derivatives of the polysaccharides*, the modified cellulose products are technically of high importance. As examples of such cellulose acetate, cellulose

nitrate, ethyl-cellulose, methyl-cellulose and benzyl-cellulose may be mentioned. The ground molecule of cellulose having 3 OH-groups, it can be esterified or etherified at 1, 2 or 3 places at will, leading to a great variety of products. In Fig. 5 cellulose triacetate is shown.
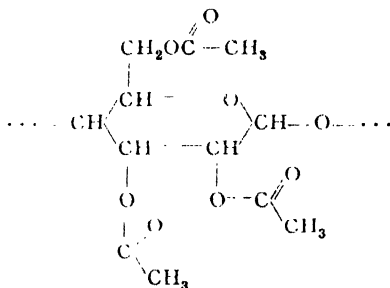
Fig. 5.   Cellulose triacetate.

2.   *Proteins* [1]

The proteins are characterised by containing main chains, built up from the structural unit:
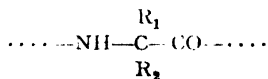
Fig. 6. Structural unit of proteins.

Here R_1 and R_2 in the simplest case are H-atoms only, but they may be extended to complicated groups, which are referred to as side-chains.

a   *The main chain*

Proteins can be divided into two series, namely:

a   The *simplest proteins*, which are built up from amino-acids only;

b   The *complex proteins* or *proteides*, which contain another component, additional to the amino-acids, the so-called prosthetic group.

Among these may be mentioned:
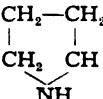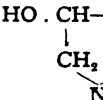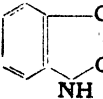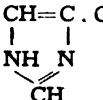
the glucoproteins, containing as prosthetic group: carbohydrates (ovalbumin, mucin)

the phosphorproteins,   ,,   ,,   ,,   ,,   phosphoric acid (casein)

the nucleoproteins,   ,,   ,,   ,,   ,,   nucleic acid

the chromoproteins,   ,,   Fe or Cu (blood pigments).

By hydrolysis [2] with dilute acids the simple proteins can be split into their

---

[1] Cf. *Symposium on Proteins*, Amsterdam 1938, published in English: *Chem. Weekbl.* (1939); D. JORDAN LLOYD and A. SHORE, *Chemistry of the proteins*, London 1938; A. FREY-WYSSLING, *Submikroskopische Morphologie des Protoplasmas*, Berlin 1938.

[2] A number of proteins are only stable in a limited pH-range. Outside that range they dissociate into smaller, well defined-units.

TABLE 3

THE NATURAL AMINO-ACIDS

	Name	Formula	pH isoelectric point
mono-amino mono-carboxylic acid	Glycine or Glycocoll	$CH_2NH_2 . COOH$	6.1
	Alanine	$CH_3 . CH NH_2 . COOH$	6.0
	Valine *)	$(CH_3)_2 CH . CH NH_2 . COOH$	
	Norleucine	$CH_3 . (CH_2)_3 . CH NH_2 . COOH$	
	Isoleucine *)	$\begin{matrix}CH_3 \\ C_2H_5\end{matrix}>CH . CH NH_2 . COOH$	
	Leucine *)	$(CH_3)_2CH . CH_2 . CH NH_2 . COOH$	6.0
	Phenylalanine *)	$\langle\rangle-HC_2 . CH NH_2 . COOH$	5.4
mono-amino dicarboxylic acid	Aspartic acid	$HOOC . CH_2 . CH NH_2 . COOH$	2.8
	Glutamic acid	$HOOC . (CH_2)_2 . CH NH_2 . COOH$	3.2
	Hydroxyglutamic acid	$HOOC . CH_2 . CHOH . CH NH_2 . COOH$	
diamino-mono-carboxylic acid	Lysine *)	$CH_2NH_2 . (CH_2)_3 . CH NH_2 . COOH$	9.9
	Arginine	$NH_2 . C=NH . NH . (CH_2)_3 . CH NH_2 . COOH$	9.0
	Citrulline	$NH_2 . CO . NH . (CH_2)_3 . CH NH_2 . COOH$	
	Ornithine	$NH_2 . (CH_2)_3 . CH NH_2 . COOH$	
sulphur containing amino-acids	Cysteine	$HS . CH_2 . CH NH_2 . COOH$	
	Cystine	$\begin{matrix}S . CH_2 . CH NH_2 . COOH \\ | \\ S . CH_2 . CH NH_2 . COOH\end{matrix}$	4.1
	Methionine *)	$CH_2S(CH_3) . CH_2 . CH NH_2.COOH$	
oxy-amino-acids	Serine	$CH_2OH . CH NH_2 . COOH$	5.7
	Tyrosine	$HO-\langle\rangle-CH_2 . CHNH_2 . COOH$	5.7
	Threonine *)	$CH_3 . \overline{CH}OH . CH NH_2 . COOH$	
hetero cyclic amino-acids	Proline	$\begin{matrix}CH_2-CH_2 \\ | \quad | \\ CH_2 \quad CH . COOH \\ \diagdown_{NH}\diagup\end{matrix}$	6.3
	Hydroxyproline	$\begin{matrix}HO . CH-CH_2 \\ | \quad | \\ CH_2 \quad CH . COOH \\ \diagdown_{NH}\diagup\end{matrix}$	
	Tryptophan	$\begin{matrix} -C . CH_2 . CH NH_2 . COOH \\ \diagdown CH \\ NH \end{matrix}$	5.9
	Histidine *)	$\begin{matrix}CH=C . CH_2 . CH NH_2 . COOH \\ | \quad | \\ NH \quad N \\ \diagdown CH\diagup\end{matrix}$	7.2

*) Cannot be synthetised by the animal body.

TABLE 4

MOLECULAR CONSTANTS OF PROTEINS

S_{20} = sedimentation constant in units of 10^{-13} reduced to water at 20°C.

D_{20} = diffusion constant in units of 10^{-7} reduced to water at 20°C.

M_s = molecular weight computed from sedimentation velocity and diffusion measurements.

M_e = molecular weight computed from sedimentation equilibrium measurements.

$M_{calc.}$ = molecular weight calculated from the rule of simple multiples.

f/f_0 = ratio of experimentally determined molar frictional constant to molar frictional constant calculated for a spherical particle of the same mass.

$\dfrac{du}{d\,pH_0}$ = slope of mobility curve in the vicinity of the isoelectric point.

Protein	S_{20}	D_{20}	M_s	M_e	$M_{calc.}$	f/f_0	pH Iso-electric Point	$\dfrac{du}{d\,pH_0}$ 10^5
Erythrocruorin (Lampetra)	1.87	10.65	17100	19100	$17600 = \frac{1}{2} . 35200$	1.2	5.60	3.2
Lactalbumin α	1.9	10.6	17500	. . . .		1.2	5.12	6.7
Cytochrome C	1.89	10.13	15600	. . . .		1.3	9.7	. . .
Myoglobin	2.04	11.25	17200	17500		1.1	7.0	7.0
Gliadin	2.00	6.72	26000	. . . .		1.6	. . .	. . .
Hordein	2.0	6.5	27000	. . . .		. . .	. . .	. . .
Zein	1.9	4.0	35000	. . . .	35200	. . .	. . .	. . .
Erythrocruorin (Arca)	3.46	. . .	. . . .	33600		1.0	. . .	. . .
Erythrocruorin (Chrironomus)	2.00	. . .	. . . .	31400		1.6	5.40	3.6
Lactoglobulin	3.12	7.27	41800	37900		1.2	5.19	11.9
Pepsin	3.3	9.00	35500	39200		1.1	. . .	. . .
Insulin	3.47	8.20	40900	35100		1.1	. . .	. . .
Bence-Jones α	3.55	. . .	. . . .	35000		1.0	5.20	5.8
Bence-Jones β	2.85	7.33	37700	. . . .		1.3	5.46	3.5
Egg albumin	3.55	7.76	43800	40500		1.1	4.55	10.4
CO-hæmoglobin (horse)	4.5	6.3	69000	68000	$70400 = 2 . 35200$	1.2	6.92	7.2
CO-hæmoglobin (man)	4.5	6.9	63000	. . . .		1.2	7.09	6.4
Serum albumin (horse)	4.5	6.17	70200	66900		1.2	4.80	9.1
Yellow ferment	5.76	6.28	82800	77800		1.2	5.22	6.4
Serum globulin (horse)	7.1	4.05	167000	150000	$140800 = 4 . 35200$	1.4	= 5.1 = 6.0	
Phycocyan (Ceramium, dissociation component)	6.2	4.58	131000	146000		1.4	4.85	10.2
Phycoerythrin (Ceramium)	12.0	4.00	290000	292000	$282000 = 8 . 35200$	1.2	4.25	14.2
Phycocyan (Ceramium, main component) .	11.4	4.05	272000	273000		1.2	4.85	10.2
Edestin	12.8	3.93	309000	. . . .		1.2	. . .	. . .
Excelsin	13.3	4.26	294000	. . . .		1.1	. . .	. . .
Amandin	12.5	3.62	329000	. . . .		1.3	. . .	. . .
Erythrocruorin (Daphnia)	16.3	. . .	. . . .	. . . .	$422000 = 12 . 35200$	. . .	. . .	. . .
Hæmocyanin (Pamdalus)	17.4	. . .	. . . .	397000		1.1	. . .	. . .
Hæmocyanin (Palinurus)	16.4	3.4	446000	447000		1.2	. . .	. . .
Hæmocyanin (Helix pomatia, dissociation component)	12.1	2.23	503000	. . . .		1.5	5.05	8.1
Hæmocyanin (Busycon, dissociation component)	13.5	3.29	379600	. . . .		1.4	4.49	10.7
Hæmocyanin (Eledone, dissociation component)	10.6	2.25	440000	. . . .		1.9	4.6	14
Thryroglobulin	19.2	2.65	628000	650000		1.5	4.58	11
Hæmocyanin (Nephrops)	24.5	2.79	820000	. . . .	$845000 = 24 . 35200$	1.2	4.64	13.3
Hæmocyanin (Homarus)	22.6	2.78	752000	803000		1.3	4.95	18
Hæmocyanin (Helix pomatia, dissociation component)	16.0	1.82	814000	797000		1.9	5.05	8.1
Hæmocyanin (Helix nemoralis, dissociation component)	16.6	1.92	799000	. . . .		1.8	4.63	11.4
Erythrocruorin (Planorbis)	33.7	1.96	1636000	1530900	$1690000 = 48 . 35200$	1.4	4.77	10.6
Hæmocyanin (Calocaris)	34.0	. . .	. . . .	1329000		1.2	. . .	. . .
Hæmocyanin (Octopus)	49.3	1.65	2785000	. . . .	$2960000 = 84 . 35200$	1.4	. . .	. . .
Hæmocyanin (Eledone)	49.1	1.64	2790000	. . . .		1.4	4.6	14
Erythrocruorin (Arenicola)	57.4	. . .	. . . .	3000000	$3380000 = 96 . 35200$	1.3	4.56	16
Chlorocruorin (Spirographis)	55.2	. . .	. . . .	. . . .		. . .	. . .	. . .
Hæmocyanin (Rossia)	56.2	1.58	3316000	. . . .		1.4	. . .	. . .
Erythrocruorin (Lumbricus)	60.9	1.81	3140000	2946000		1.2	5.28	12.6
Hæmocyanin (Helix pomatia, main component)	98.9	1.38	6630000	2680000	$6760000 = 192 . 35200$	1.2	5.05	8.1
Hæmocyanin (Busycon, main component) .	101.7	. . .	. . . .	. . . .		1.2	4.49	10.7
Hæmocyanin (Busycon, aggregation component)	130.4	. . .	. . . .	. . . .	$10140000 = 288 . 35200$	1.2	4.40	10.7

components, the amino-acids. This gives an insight into their chemical constitution, because it is assumed that nature builds up[1] its proteins from these acids as follows:

$$
\begin{matrix} R_1 & & R_1' & & R_1 & & R_1' \\ H_2N \cdot C{-}CO\,OH & + & H_2N{-}C{-}COOH & \rightarrow & H_2N{-}C{-}CO{-}NH{-}C{-}COOH & + & H_2O \\ R_2 & & R_2' & & R_2 & & R_2' \end{matrix}
$$
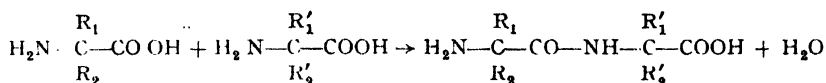
Fig. 7. Building up of proteins from amino-acids.

To give an idea of the great variety of proteins which exists, considered only from a chemical point of view, the amino-acids known to be present in proteins, are summarized in Table 3. From the colloidal point of view the situation is still more complicated because there are linear proteins (fibrin, glutin), and globular proteins also (albumin, globulin). Furthermore, there are also proteins with an interlinked net-structure (keratin).

We owe to BERGMANN[2] a far-reaching theory concerning a certain regular pattern which seems to be followed in building up the proteins from these amino-acids. Before describing this however, the magnificent results of SVEDBERG's molecular weight determinations by means of the ultracentrifuge[3] must be referred to briefly (cf. for a detailed description Chap. V, p. 131). He found that the globular proteins consist of molecules with a weight that, neglecting small deviations, is a multiple of 17600. Thus, insulin (mol.wt. = 35100) is built up of 2, and hæmoglobin (mol.wt. = 68000) of 4 such "SVEDBERG units", this number increasing to 576 (mol.wt. over 10[7]), as is demonstrated in Table 4.

Since each SVEDBERG unit contains about 294 residues, one can say, that the polymerisation degree of the SVEDBERG unit is 294, the lower proteins thus being mesocolloidal. Above a molecular weight of 35200 however, they are already getting

TABLE 5

PROBABLE CONSTITUTION OF HAEMOGLOBIN

lysine	$\ldots - L - 15x - L - 15x - L - \ldots$
hystidine	$\ldots - H - 17x - H - 17x - H - \ldots$
glutamic acid	$\ldots - Gl - 35x - Gl - 35x - Gl - \ldots$
asparagine	$\ldots - A - 17x - A - 17x - A - \ldots$
tyrosine	$\ldots - T - 47x - T - 47x - T - \ldots$
proline	$\ldots - Pr - 47x - Pr - 47x - Pr - \ldots$
arginine	$\ldots - Ar - 47x - Ar - 47x - Ar - \ldots$
cysteine	$\ldots - C - 191x - C - 191x - C - \ldots$

into the eucolloidal range. We showed on p. 24, that in our classification even the largest protein molecules are still to be considered as colloidal.

Now coming to the far-reaching conception of BERGMANN, he assumes that the polymerisation degree of the SVEDBERG unit should be exactly $288 = 2^5 \cdot 3^2$, instead of 294. Generalizing this, the total number of residues in the larger protein molecules

[1] There are probably however, in proteins other types of bonds apart from the polypeptide bond. An argument in favour of this is that the enzyme pepsin can destroy proteins but not polypeptides.
[2] M. BERGMANN and C. NIEMANN, *J. Biol. Chem.*, 118 (1937) 301; *Ibid; Science*, 86 (1937) 187.
[3] T. SVEDBERG, *Nature*, 139 (1937) 1046, 1051; T. SVEDBERG and K. PETERSEN, *Die Ultracentrifuge*, Dresden 1940.

should be expressed by $N_t = 2^n \cdot 3^m$, whereas the number of each individual amino-acid should be $N_i = 2^{n'} \cdot 3^{m'}$. The frequency with which each residue occurs in a molecule is then $\frac{N_t}{N_i}$, which can be expressed by $2^{n''} \cdot 3^{m''}$. In these formulae n, m, n', m', n'' and m'' are either zero or a whole number, and obviously, $n = n' + n''$, $m = m' + m''$ and $N_t = (N_i)_a + (N_i)_b \ldots + (N_i)_x$, where a, b ... x indicate the various amino-acids. Each amino-acid therefore seems to repeat itself in the protein molecule according to a certain pattern. For example, Table 5 represents the situation for haemoglobin, as far as it is known at the present time.

TABLE 6

SOME EXAMPLES OF SIDE-GROUPS R IN POLYPEPTIDE CHAINS

Type	side group	Type	side group
acid	$\diagup$ NH $\diagdown$ CH—CH$_2$—COOH CO $\diagdown$ asparagic acid	basic	$\diagup$ NH $\diagdown$ CH—CH$_2$—CH$_2$—NH$_2$ CO $\diagdown$ ornithine
	$\diagup$ NH $\diagdown$ CH—CH$_2$—CH$_2$·COOH CO $\diagdown$ glutamic acid		$\diagup$ NH $\diagdown$ CH–CH$_2$–CH$_2$–CH$_2$–NH–C$\diagup$ NH $\diagdown$ NH$_2$ CO $\diagdown$ arginine (valine + guanidine)
hydrophobic	$\diagup$ NH $\diagdown$ CH—CH$_2$—CH $\diagup$ CH$_3$ $\diagdown$ CH$_3$ CO $\diagdown$ leucine	hydrophilic	$\diagup$ NH $\diagdown$ CH—CH$_2$—OH CO $\diagdown$ serine
	$\diagup$ NH $\diagdown$ CH—CH$_2$—⬡ CO $\diagdown$ phenylalanine		$\diagup$ NH $\diagdown$ CH—CH$_2$—⬡—OH CO $\diagdown$ tyrosine
sulphur containing	$\diagup$ NH $\diagdown$ CH—CH$_2$—SH CO $\diagdown$ cystine	possible chainend	CH$_2$ $\diagup$ $\diagdown$ NH CH$_2$ $\diagdown$ CH$_3$ $\diagdown$ CH $\diagup$ CO $\diagdown$ proline

This means, that hæmoglobin would consist of at least 8 amino-acid residues of which lysine occupies every 16th place in the chain, histidine every 18th place, etc. When later the patterns of all constituent amino-acids are known, these must fit together in such a way, that no place x will remain unknown.

b *The side chains*

In many books one can find, that the amphoteric character of the proteins is due to the presence of acid (COOH) and basic (NH_2) groups in the amino-acids. This, however, cannot be right as these groups disappear, according to Fig. 7, p. 32, during the formation of the protein. The causes of the amphoteric properties therefore must be sought in the side chains, and from Table 6, it may be seen that COOH- and NH_2-groups are found in certain chains. This is one of the reasons, that the *side-chains* are perhaps even more important, from a colloidal point of view, than the polypeptide main chains. In Table 6 some examples of side-groups are given in order to show the great variety of types which exist.

We also find here side-groups which are responsible for hydrophobic properties, because they bear CH_3 or C_6H_5 groups at their end. On the other hand, OH-groups, as in serine, may lead to hydrophilic proteins. In addition, the possibilities of bridge-formation between the polypeptide chains are dependent on the chemical constitution of the side groups.

In Fig. 8 four cases of fundamental importance to our subject are pictured. The first is the VAN DER WAALS's bond between hydrophobic groups (attraction between CH_3-groups), which can only be loosened by hydrophobic solvents. The second is the dipole attraction (e.g., between 2 OH-groups), which can only be loosened by polar groups like OH, with the result that water causes swelling. This swelling is greatly influenced by salts, as is indicated at the right hand of the figure. The third and fourth types contain *primary bonds* (e.g., the sulphur-sulphur bond or the bonds in the main chain) of high energy content, which may be homopolar or

1^0. VAN DER WAALS's attraction between hydrophobic groups

2^0. dipole attraction between hydrophilic groups

3^0. heteropolar bond (salt-formation)
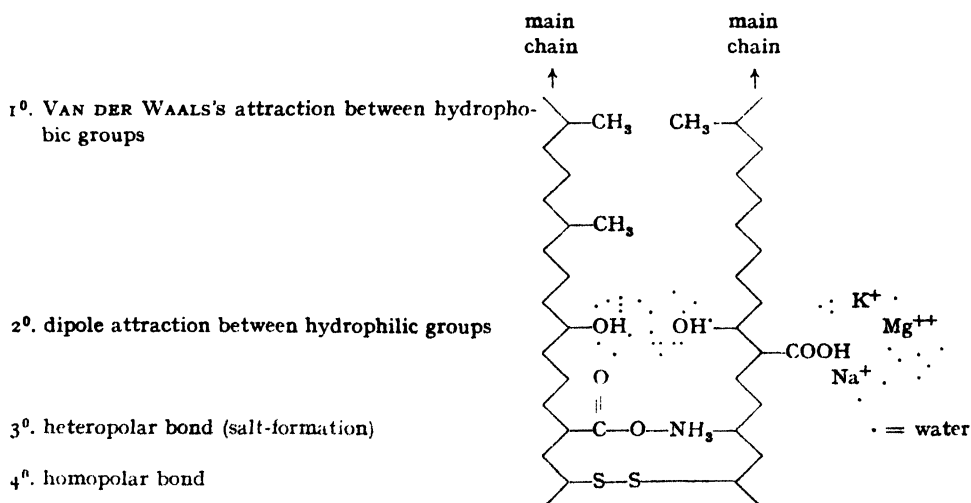
4^n. homopolar bond

Fig. 8. Four possibilities of bridge-formation between protein chains.

heteropolar. The heteropolar bond can still be loosened by water as a consequence of the high solvation energy of the latter. The homopolar bond cannot as a rule be loosened by solvents, but only by heating or other such drastic means.

Another factor, in which the constitution of the side-chains plays a rôle is the crystallising power. When they are of very unequal length, a regular pattern cannot be formed and crystallisation will not be possible.

Among the *modified proteins*, products hardened by means of formaldehyde may be mentioned. Such hardening processes are being increasingly applied in industry. The conversion of casein into artificial horn has been known for a long time; the reaction may be of the type shown in Fig. 9, where two protein chains are connected[1] by means of a CH_2-group. It also seems likely that in the hardening of casein-wool (from milk) similar reactions play a part.

$$\cdots\text{CO}-\text{NH}-\overset{\displaystyle R_1}{\underset{\displaystyle R_2}{\text{C}}}-\cdots + CH_2O \rightarrow$$

$$\begin{array}{c} R_1 \\ \cdots -\text{CO}-\text{N}-\text{C}-\cdots \\ |\quad R_2 \\ \text{CH}_2 \\ |\quad R_3 \\ \cdots -\text{CO}-\text{N}-\text{C}-\cdots \\ R_4 \end{array}$$
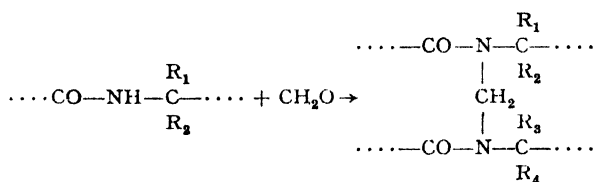
Fig. 9. Possible modification (interlinking) of proteins by means of formaldehyde

3. *Rubber*[2] being polyisoprene, its formula is represented by

$$\cdots(-CH_2-CH=\overset{\displaystyle }{\underset{\displaystyle CH_3}{C}}-CH_2)_n \cdots$$
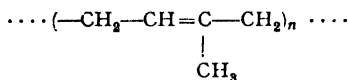
Fig. 10. Natural Rubber.

It will appear later that gutta percha, having the same molecular formula $(C_5H_8)_n$, differs in so far as it is a stereoisomer, leading to entirely different properties (see p. 663)

It is possible to make a whole series of *derivatives of rubber*. First one can add atoms to the double bond; the products, obtained with Cl_2 and HCl, being well-known. Apart from this however, substitution also occurs so that chlorinated rubber may be represented by

$$\begin{array}{c} CH_3 \\ | \\ -CH-C-CH-CH- \\ |\quad |\quad |\quad | \\ Cl\quad Cl\quad Cl\quad Cl \end{array}$$
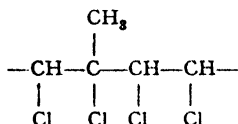
Fig. 11. Chlorinated rubber.

The action of sulphur on rubber (vulcanisation) has a similar interlinking effect as was described for proteins and formaldehyde; we will consider this process more fully on p. 675.

---

[1] Cf., for the technical importance of such "interlinking" reactions p. 673.
[2] Cf. K. H. Meyer, *l. c.* on p. 26.

4. *Drying oils* are compounds of three molecules of fatty acid, attached to one molecule of glycerine. The technically important fatty acids involved are represented in Table 7 and vary by their degree of unsaturation or by the presence of an OH-group (ricinoleic acid) or a carbonyl-group (couepinic acid). Of great additional importance is whether the double bonds are conjugated (wood-oil) and whether the chain has an unbroken row of seven (CH_2)-groups. On heating, these oils gelatinise as a consequence of net formation, which is probably due[1] to an addition between

TABLE 7

UNSATURATED FATTY ACIDS

$CH_3 - (CH_2)_7 - CH = CH(CH_2)_7 COOH$	oleic acid
$CH_3 - (CH_2)_4 CH = CH - CH_2 - CH = CH - (CH_2)_7 COOH$	linoleic acid
$CH_3 - (CH_2 - CH = CH)_3 - (CH_2)_7 COOH$	linoleic acid
$CH_3 - (CH_2)_5 = CHOH - CH_2 - CH = CH - (CH_2)_7 COOH$	ricinoleic acid
$CH_3 - (CH_2)_3 - (CH = CH)_3 - (CH_2)_7 COOH$	eleostearic acid (wood oil)
$CH_3 - (CH_2)_3 - (CH = CH)_3 - (CH_2)_4 - CO - (CH_2)_2 COOH$	couepinic acid (oiticica oil)

conjugated and ordinary double bonds, leading to ring formations as pictured in Fig. 12. Oxygen can also play an interlinking rôle and all such reactions lead to highly important technical consequences (see p. 673).
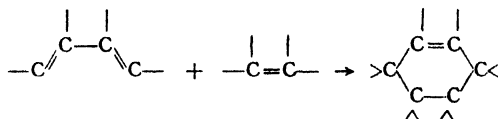
Fig. 12.   Ring formation during polymerisation of drying oils.

Furthermore, drying oils can be combined with all kinds of other substances; the results obtained with synthetic resins are of extreme importance for varnish industry.

5. *Natural Resins*

Some natural resins possess molecules of colloidal size. Others become colloidal either by polymerisation on heating, or by the action of foreign substances, which bring about an interlinking of the molecules by bridges.

The most important active natural resins are shellac, colophony, damar and copal, which contain active groups like COOH, OH or double bonds.

c *Entirely synthetic macromolecules*

On p. 40 a distinction is made between macromolecules formed by polymerisation and by condensation. In the first case growth takes place without, and in the second

---

[1] C. P. A. KAPPELMEIER, *Paint, Oil, Chemical Rev.*, 100 No. 1—5, 2—9 and 4—9 (1938).

TABLE 8

SYNTHETIC POLYMERS

a. *Based on ethylene*

$$\cdots - \left[CH_2-CH_2 \right]_x - \cdots \quad \text{Polyethylene}$$

$$\cdots - \left[\begin{array}{c} CH-CH_2 \\ | \\ COOH \end{array} \right]_x - \cdots \quad \text{Polyacrylic acid}$$

$$\cdots - \left[\begin{array}{c} CH-CH_2 \\ \bigcirc \end{array} \right]_x - \cdots \quad \text{Polystyrene}$$

$$\cdots - \left[\begin{array}{c} CH-CH_2 \\ | \\ CO-OC_2H_5 \end{array} \right]_x - \cdots \text{Polyacrylic ethyl ester}$$

$$\cdots - \left[\begin{array}{c} CH-CH_2 \\ | \\ Cl \end{array} \right]_x - \cdots \quad \text{Polyvinyl chloride}$$

$$\cdots - \left[\begin{array}{c} CH-CH_2 \\ | \\ OCO \\ | \\ CH_3 \end{array} \right]_x - \cdots \quad \text{Polyvinyl acetate}$$

b. *Based on isobutylene*
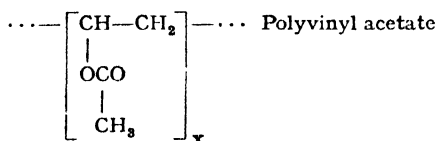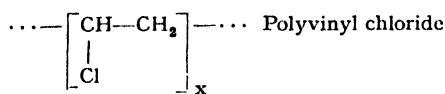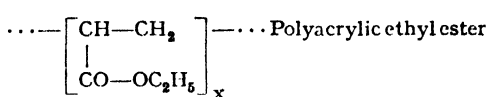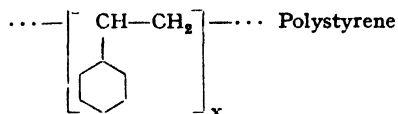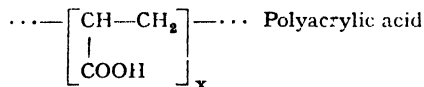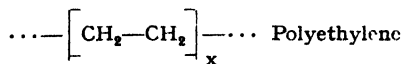
$$\cdots - CH_2 \left[\begin{array}{c} CH_3 \\ | \\ -C-CH_2 \\ | \\ CH_3 \end{array} \right]_x - \cdots \quad \text{Polyisobutylene}$$

c. *Based on butadiene*

$$\cdots \left[\begin{array}{c} H_2 \; Cl \; H \; H_2 \\ -C-C=C-C \end{array} \right]_x - \cdots \quad \text{Chloroprene (Neoprene)}$$

$$\cdots \left[\begin{array}{c} H_2 \; H \; H \; H_2 \; H_2 \quad H \\ -C-C=C-C-C-\!-C \\ \qquad\qquad\qquad\qquad | \\ \qquad\qquad\qquad\qquad \bigcirc \end{array} \right]_x - \cdots \quad \text{Co-polymer G.R.-S (Government rubber S; German: Buna S)}$$

butadiene-   styrene-
unit   unit

$$\cdots \left[\begin{array}{c} H_2 \; H \; H \; H_2 \; H_2 \; H \\ -C-C=C-C-C-\; C \\ \qquad\qquad\qquad\qquad | \\ \qquad\qquad\qquad\qquad CN \end{array} \right]_x - \cdots \quad \text{Co-polymer Perbunan or Buna N}$$

butadiene-   vinyl
unit   cyanide-
           unit
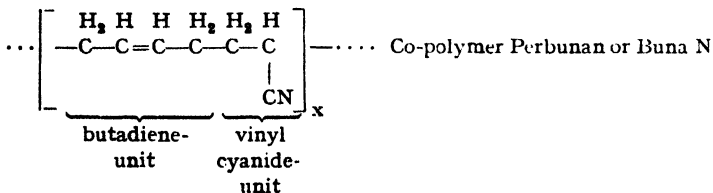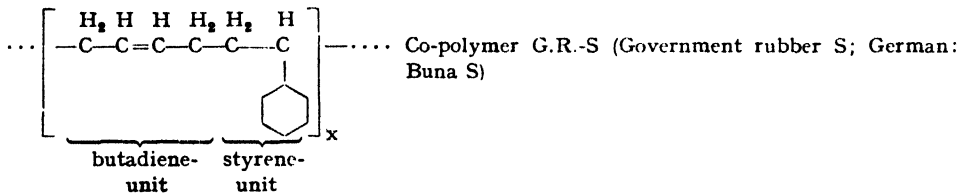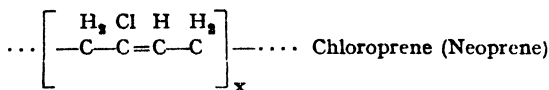
case, accompanied by a formation of foreign products (e. g., H_2O). A vast number of examples[1] of both kinds are known. We will confine ourselves to those of special importance from a colloidal point of view.

For the polymerisates there are three main groups, according to Table 8, based on ethylene, isobutylene and butadiene respectively. The first group as a rule contains the so-called thermoplastics, being hard products which can be weakened by heating. The two other groups contain rubbery materials.

Among the *condensates* (Table 9) are found the thermosetting synthetic resins which, in contrast to the thermoplastics, become harder[2] on heating. They are all products with globular molecules, among which phenol-formaldehyde and urea-formaldehyde are well known. In addition there are the so called alkyd-resins (*alcoholacid*), obtained by the condensation of polyvalent alcohols and acids. Among these last products, playing an outstanding part in film formation in the varnish industry, the esters of glycerine with the dicarboxylic acids, maleic acid and phthalic acid, may be mentioned. They are usually combined with drying oils, but it is also possible to combine some of them with natural resins. Altogether this has led to a new branch of industry, based entirely on the interlinking of macromolecules.

At the end of Table 9 we find under d) the rubbery product polyethylene tetrasulphide and under e) the so-called synthetic polyamides (Nylon), the latter being condensation products between diamines and dicarboxylic acids. In contrast to the others, the types d) and e) are linear polymers.

d Inorganic macromolecules[3]

After the fundamental remarks in § 1 about the existence of macromolecules, it is clear that the group of inorganic representatives must be treated separately.

Linear polymers. Among these *asbestos* is one of the best known examples. As however this material offers no important viewpoints for colloidal considerations, we will not consider it further.

More interesting is *sulphur*, obtained by quick cooling after being kept for some time in the molten state at 250° C. The material then becomes rubbery and possesses chain-molecules of the following constitution:
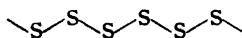
Fig. 13.   Constitution of the chain-molecules in rubbery sulphur.

Selenium can form chain-molecules at 70—73° C and then it too shows rubbery properties.

---

[1] For a summary see: C. ELLIS, *The Chemistry of Synthetic Resins*, New York 1935; T. HEDLEY BARRY, *Natural Varnish Resins*, London 1932.

[2] In some cases after being softened previously for a short while.

[3] A summary is given by K. H. MEYER, *Chemistry and Industry*, 57 (1938) 439; K. H. MEYER, *Natural and Synthetic High Polymers*, New York 1942.

TABLE 9
SYNTHETIC CONDENSATION PRODUCTS

a)

Phenol   formaldehyde   Trihydroxy-trimethyl-phenol

Phenol-formaldehyde resin ("Bakelite")

b)

Urea   formaldehyde

Urea-formaldehyde resin ("Pollopas", "Beetle")

c)

Glycerine   phthalic acid   Glyptal resin

d)

Dichlor-ethylene   Na poly-sulphide   Polyethylene tetra-sulphide ("Thiokol")

e) $H_2N(CH_2)_p NH_2 + HOOC(CH_2)_q COOH \rightarrow$
    diamine   dicarbonic acid

$\cdots -HN(CH_2)_p[NHCO(CH_2)_q CONH(CH_2)_p]_x NHCO(CH_2)_q -\cdots$
synthetic polyamide ("Nylon")

Poly-phosphorus chloronitride becomes rubbery when heated to 300° C and then chain-molecules of the type shown in Fig. 13a are formed.

$$\begin{array}{ccc} Cl & Cl & Cl \\ -P{=}N{-}P{=}N{-}P{=}N{-} \\ Cl & Cl & Cl \end{array}$$
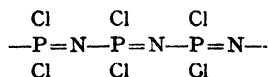
Fig. 13a. Chain-molecule of poly-phosphorus chloronitride.
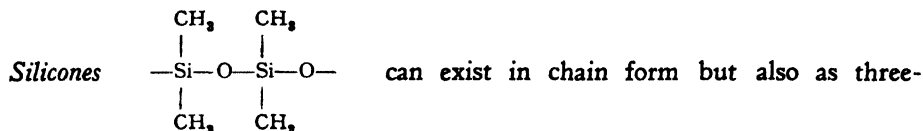
Silicones
$$\begin{array}{cc} CH_3 & CH_3 \\ | & | \\ -Si{-}O{-}Si{-}O{-} \\ | & | \\ CH_3 & CH_3 \end{array}$$
can exist in chain form but also as three-

dimensional molecules. In the first case it shows some analogy with crude rubber and in the second case with vulcanised rubber.

e *Two-dimensional polymers*

Graphite is one of the best known examples of this type; it has already been treated fully in Volume I of this book.

Among the *silicates* there are a lot of representatives, with planar molecules. As examples, one may mention mica, clay and talcum which have all been dealt with in Volume I.

§ 6   THE BUILDING UP OF MACROMOLECULES IN NATURE AND IN PRACTICE; GEL AND ISOGEL FORMATION

Since much more is known about the mechanism of the building up of macro-molecules in synthetic processes than in nature, we will first consider the results obtained in the laboratory.

A *Building up in practice*[1]

There are two different ways of building up macromolecules namely *polymerisation* and *condensation*.

In polymerisation processes monomeric molecules attach themselves to each other without the formation of any new reaction product apart from the macro-molecules. The polymerisation of vinyl chloride in Fig. 14 can be considered as representative of this type; it is a simple addition.

$$\begin{array}{ccc} CH_2{=}CH & \rightarrow & \cdots{-}CH_2{-}CH{-}CH_2{-}CH{-}\cdots \\ | & & | & | \\ Cl & & Cl & Cl \\ \text{monomer} & & \text{polymer} \end{array}$$
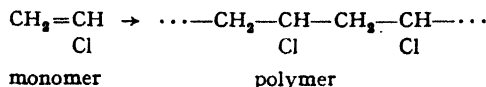
Fig. 14. Polymerisation of vinyl chloride.

[1] For comprehensive literature, c.f., E. BURK, H. S. THOMPSON, A. J. WEITH, I. WILLIAMS, *Poly-merization*, New York 1937; R. HOUWINK, *Elastomers and Plastomers*, Amsterdam 1949, Vol. I.

In contrast to this stands the type shown in Fig. 15, in which the reaction between glycol and maleic acid is pictured and which involves the formation of H_2O.

$$
\begin{array}{cccccc}
\text{OH} & \text{OH} & \text{COOH} & \text{COOH} & \text{OH} & \text{O—C} \overset{O}{\diagup} \quad \text{C} \overset{O}{\diagdown} \text{O} \\
| & | & | & | & | & | \quad | \quad | \quad | \quad | \\
\text{CH}_2\text{—CH}_2 & + & \text{CH}\!\!=\!\!\!=\!\!\text{CH} & \rightarrow & \text{CH}_2\text{—CH}_2 & \text{CH}\!=\!\text{CH} \quad \text{CH}_2 \quad + H_2O
\end{array}
$$

glycol   maleic acid
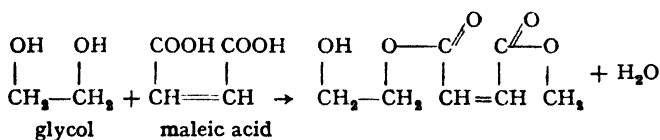
Fig. 15. Alkyd resin building.

In this case therefore an *equilibrium* is produced, so that the reaction can only be completed by removing the water. Polymerisation reactions however are chain reactions, which can be brought to completion without the aid of such artificial means.

As the mechanism of condensation reactions shows most connection with classical low-molecular chemistry, we will consider such reactions first.

a Condensation reactions

Growth takes place in similar successive stages because reactive groups are always available at the ends of the molecules. The reaction equilibrium [1] is expressed by

$$\frac{C\ \text{ester} \cdot C\ \text{water}}{C\ \text{acid} \cdot C\ \text{alcohol}} = K \tag{1}$$

where C denotes the concentration. At any moment a whole series of molecular sizes, are present according to the formula:

$$F_n = q^{n-1}(1-q) \tag{2}$$

where F_n = molecular fraction of the polymerisation degree n; q = relative number of reactive groups used up in the condensation.

It is clear that the colloidal behaviour of polymers will be largely dependent on the distribution of the molecular sizes. In Fig. 16 the change of this distribution with increasing q is shown for an esterification reaction.

It can be concluded from these curves that even when the reaction has proceeded to the extent of 50% there are still no large molecules formed. These only appear at the end, but there are still a certain number of small molecules present even when 99% of the esterification has been completed, although their number diminishes gradually. In practice, where the last traces of water are difficult to remove, one
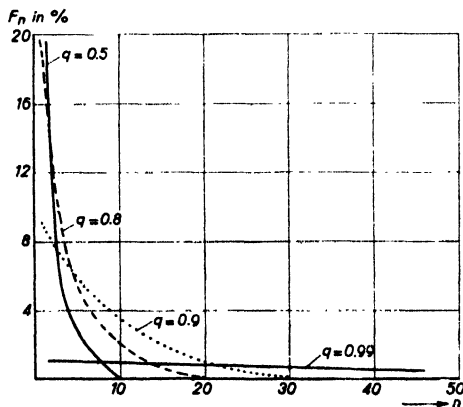
Fig. 16. Change of distribution of molar fractions F_n with increasing degree of esterification q.

---

[1] Cf., for original studies about the kinetics of condensation reactions: R. H. KIENLE and A. C. HOVEY, J. Am. Chem. Soc., 52 (1930) 3636; J. Soc. Chem. Ind., 55 (1936) 229 T; P. J. FLORY J. Am. Chem. Soc., 58 (1936) 1877; 59 (1937) 241; H. MARK and H. DOSTAL, Trans. Faraday Soc., 32 (1936) 54.

has always to deal with a mixture of small and large molecules, which are usually for the greater part soluble in each other, leading to the conditions for gel formation. These will be partly *isogels*, because the small and large polymer molecules have the same chemical constitution. But apart from this, water molecules create a condition for *heterogel formation*, so that one has both types to deal with, in contrast to the situation we will meet when discussing polymerisation reactions.

b   *Polymerisation reactions*

It is generally assumed at the present time that polymerisations take place by three different steps, as follows:

a   The *start reaction*[1], in which the monomer molecule is activated to form a radical according to Fig. 17.

$$CH{=}CH_2 \longrightarrow -CH-CH_2- \longrightarrow \cdots-CH-CH_2-CH-CH_2-CH-CH_2-\cdots$$

| | | |
| Cl | Cl | Cl   Cl   Cl |

start   chain
reaction   growth

monomer   monomer   macro-radical
molecule   radical

$$\longrightarrow \quad CH-CH_2-CH-CH_2--CH{=}CH_2$$

final   Cl   Cl   Cl
reaction   macromolecule

Fig. 17. Stages in the polymerisation of vinyl chloride.

b   The *chain growth*, in which a great number of the activated monomer radicals are attached to each other, forming the polymer chain which however is still a radical (macro-radical).

c   The *final reaction*, by which the macro-radical is converted into a macro-molecule, losing its reactive power.

To explain these various reactions, one may consider that the formation of the monomer radical is due to the fact that now and then a monomer molecule is brought into a state of very high energy, e.g., as a consequence of a specially active thermal collision, the absorption of a light quantum, or some other process. This may lead to an "opening" of the double bond as indicated in the figure. For this type of activation, between 20 and 50 kcal/mol are usually necessary. Such high values lead to the fact that even at 100° C only 10^{-12} to 10^{-25} of the collisions are of sufficiently high energy content to result in the initial reaction, and for this reason the use of catalysts at this stage results in a greater efficiency by lowering the energy barrier.

To explain the chain growth reaction, it can be supposed that at the moment of collision so much energy is transferred to the second monomer molecule that this

---

[1] Also denoted as activation reaction or nuclei formation.

also opens its double bond and at that moment the electrons can unite into a pair leading to the formation of a homopolar bond.

As an energy transfer from a monomer radical to a macro-radical is all that occurs, the heat of activation of this type of reaction is much lower, namely of the order of 3 to 6 kcal/mol, leading to a reaction velocity which is from 10^7 to 10^{12} times greater than that of the start reaction at the same temperature. The situation therefore is often that chain growth cannot occur, because no monomer radicals have yet been formed, so that the velocity of the whole polymerisation is governed by the velocity of the start reaction. As soon as a monomer radical has been formed however, the chain growth as a rule will start immediately.

Various causes for this final reaction may be considered. For example, one can think of the migration of a H-atom along the chain, of a reaction between two macro-radicals, or of a ring formation between the chain ends themselves. Another cause may be that, when at the moment of collision the radical is not in the most suitable position on steric grounds, the energy transfer does not take place. This steric factor, which is negligible for small radicals, becomes greater as the molecular sizes increase[1] and therefore it will be especially unfavourable with macromolecules. This automatically brings about a limit to the mean molecular length.

The resulting polymerisation velocity will of course depend on the reaction constants K_1, K_2, and K_3 of the three reactions involved. When these are of the same order of magnitude the situation becomes very complicated. Other complications may arise from the use of catalysts[2] inhibitors[3], the occurrence of bimolecular reactions[4], etc., some of which have been pictured in
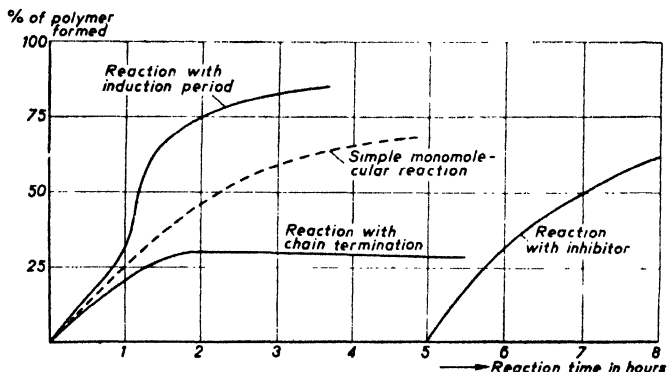
Fig. 18. Various types of polymerisation reactions.

Fig. 18. There are however cases, as already mentioned above, where the situation is extremely simple, namely when the initial reaction is so slow in comparison with the chain growth and the final reaction that the final result is dominated by the formation process of the monomeric radical.

Whatever the type of reaction, a series of polymer molecules of different length

---

[1] With addition reactions of large organic molecules a decrease of the reaction probability to 10^{-11} has been observed; cf., C. E. A. BANN, *Trans. Faraday, Soc.*, 32 (1936) 178; EVANS and M. POLANYI, *Trans Faraday Soc.*, 31 (1935) 875; H. EYRING, *J. Chem. Physics*, 3 (1935) 107; A. WASSERMANN, *J. Chem. Soc., London,* (1936) 432, 1028.

[2] Oxygen from the air often acts as a catalyst by forming a peroxide; cf., A. C. CUTHBERTSON, G. GEE, and E. K. RIDEAL, *Proc. Roy. Soc. London,* A. 170 (1939) 300.

[3] Cf., the action of hydroquinone as an inhibitor; H. S. TAYLOR and A. A. VERNON, *J. Am. Chem. Soc.*, 53 (1931) 2527.

[4] E. g., as a result of a particularly high energy collision of two monomer molecules.

will always result, because very rarely will the three conditions for the activation, chain growth and final reaction be exactly alike for two monomeric molecules. Therefore here also, we always have to deal with *distribution curves of the molecular size, (polydispersion)*, of which examples are given in Fig. 19a and Fig. 19b, referring to the number[1]
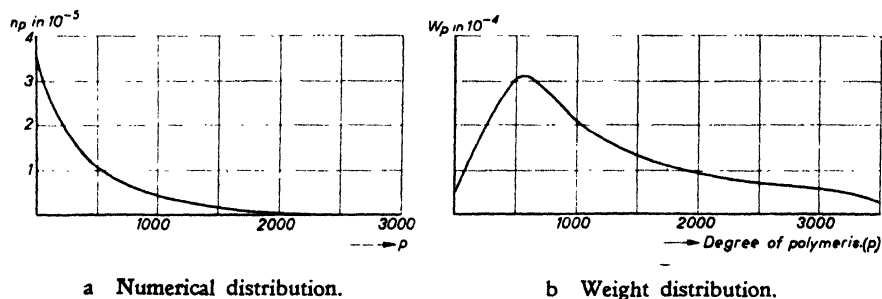
a   Numerical distribution.   b   Weight distribution.
Fig. 19.   Distribution of the various degrees of polymerisation in poly*iso*butylene[2].

and weight[3] of the particles respectively. The great difference in form between these two curves is caused by the fact that the small molecules, although present in a great number, form only a relatively small weight fraction.

As very small particles (even monomer particles, which of themselves would form a liquid) are nearly always present, having the same chemical constitution[4] as the polymer molecules, the ideal conditions are present for pure *isogel-formation*[5]. This differs therefore from the gel formation found with condensation polymers in which a mixture of isogels (the various particles of the polymer themselves) and of heterogels (water and the polymer) are present.

It will be seen later, that the plastic properties of isogels are much influenced by the amount of monomer present, as the latter forms a lubricator by means of which the macromolecules can glide over each other.

Another important technical point is the state of division in which the polymerisation reaction is being carried out. The oldest process is the so-called *block-polymerisation*, where the monomer is heated without any addition of dispersing media. It gradually hardens, becoming a solid mass. As this cannot be stirred at the end of the process, it is impossible to maintain an equal temperature throughout the mass, leading to an unhomogeneous product from the point of view of molecular size[6]. Attempts have been made to overcome this difficulty by carrying out the *polymerisation in solution*, which of course enables stirring, but has the drawback that afterwards the solvent must be evaporated. An ideal solution of all these difficulties has been found by introducing *emulsion-polymerisation*. Here the monomer is emulsified in

---

[1] N_p denotes the number of gram-molecules with a certain degree of polymerisation p, present in a weight equal to the molecular weight of the structural unit.
[2] G. V. SCHULZ, *Z. phys. Chem.*, B. 30 (1935) 379; B. 32 (1936) 27.
[3] W_p denotes the number of grams with a certain degree of polymerisation p which are present in a gram of the material.
[4] Apart from the end groups.
[5] The same is true of course for isosols; see p. 660.
[6] At a higher temperature a lower molecular weight is obtained.

water and the drops are polymerised, being coagulated later in order to form a coherent mass of polymer. It is remarkable that nature also seems to apply this emulsion-polymerisation process, discovered in technics only a few years ago (see below).

c *Building up in nature*

When considering the building up of macromolecules in Nature, we will not discuss the particular phenomenon of metabolism[1], that in the living animal body molecules are continuously in a state of transition, being broken down steadily on the one hand, and being restored simultaneously on the other hand. This phenomenon is observed not only in storage materials like fats and sugars but also in structural parts like proteins. Proof of such reactions could be obtained by means of isotopes; the changes in the body of certain fats and proteins "labelled" by means of deuterium or a nitrogen isotope have been followed for this purpose.

The reason for not discussing these phenomena is that in this book colloidal systems at a given moment are considered independently of changes occurring in the chemical constitution of the molecules.

Although very little exact information is known about the methods by which Nature builds up its macromolecules, it is probable that condensations and poly-merisations, as described in the previous paragraph, are both involved. Apart from this however, the possibility is also left open that no monomer at all is formed but that other products like CO_2 and H_2O are condensed directly into chain molecules. We will not describe in detail the abundant number of theories[2], which have been developed in these fields. Only a few examples will serve as an illustration, starting with the *condensations* as examples from classical chemistry.

In this connection, Fig. 7 on p. 32 may be referred to, where the condensation reaction between two amino-acids in the building up of proteins is shown. The next example can be found with starch. This is built up by CO_2-assimilation and here two processes seem to play a part, first a reduction and then a synthesis. Another example, in which Nature seems to use simple reactions from classical chemistry, is the building up of lignin. Although its constitution is not precisely known, it seems to contain dihydroxyphenyl-glycerine, which, as FREUDENBERG[3] suggests, condenses according to Fig. 20.

Dihydroxyphenyl-
glycerine   Condensation product
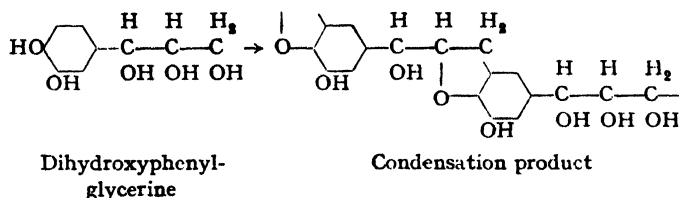
Fig. 20. Suggested reaction for the formation of lignin.

[1] See R. SCHOENHEIMER, *The Dynamic State of Body Contituents*, Cambridge (Mass.) 1946.

[2] A survey is given by A. FREY-WYSSLING, *Die Stoffausscheidung der höheren Pflanzen*, Berlin 1935.

[3] K. FREUDENBERG and W. DÜRR, *Kleines Handb. der Pflanzenanalyse*, 3 (1932) 142; K. FREUDENBERG and SOLMS, *Ber. dtsch. chem. Ges.*, 66 (1933) 262.

This polymerisation should continue at length, building up lignin three-dimensionally because an interlinking in all directions is possible.

Cases are however also known, where *polymerisations* are assumed to occur in Nature. RUZICKA[1] has developed a very general theory, according to which all terpenes might be built up from isoprene. Including further the presence of certain oxygen-containing terpenes, the formation of a great many resins (colophony), vitamins, sapogenin, carotenoids, camphor, and rubber is considered from this point of view. He assumes that this polymerisation leads either to rings or chains, according to the schemes of Fig. 21. The dotted lines indicate the places at which the isoprene units are attached to each other.
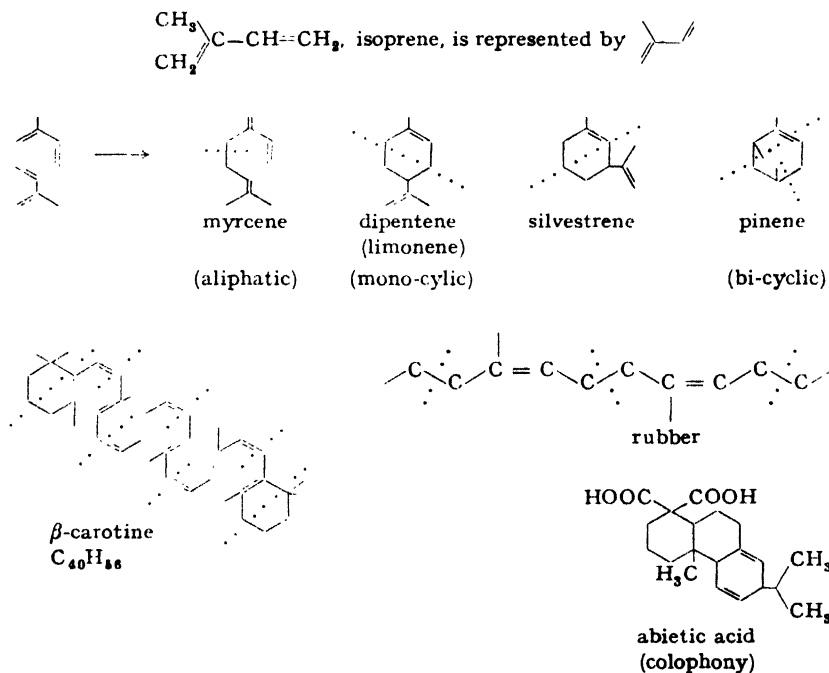
Fig. 21. Some natural polymers, built up from isoprene or oxygen containing terpenes.

Another question[2] is how and where the polymerisation takes place in the plant. There are authors, who assume that the polymerisation of isoprene to form rubber is already complete when the rubber flows out of the tree. Others [3] however, suppose the polymerisation takes place after leaving the tree, namely at the moment of coagulation [4]. In the last case the reaction would proceed very rapidly and would be an

[1] Statements of a lecture of L. RUZICKA about this subject are found in A. FREY-WYSSLING, *loc. cit.*
[2] Cf., for a survey: G. VAN ITERSON JR., *Ind. Rubber J.*, 92 (1936) 869.
[3] J. H. E. HESSELS, *Diss.*, Delft 1943.
[4] This polymerisation should be started at the surface of the latex globules by protein catalysts, which are always present there and under the further influence of light and oxygen.

emulsion-polymerisation. As already remarked above this would lead to a striking similarity between Nature and modern technique; in our opinion however, there is no much evidence for the occurrence of this type of reaction in Nature.

From experiments by STAUDINGER[1], who showed that the younger parts of plants as a rule contain shorter molecules than the older parts, it may be taken that, in the case of cellulose the molecules grow very slowly. He found for instance, the following values.

TABLE 13

POLYMERISATION DEGREE IN YOUNGER AND IN OLDER PARTS OF THE PLANT

	pol. degree
points of fir-needles (very young)	1000
leaves of oaks	1350
wood-cellulose (usual)	2000
young rye	490
leaves of rye	1000
blade of rye	1160
rye-straw (ripe)	1610

It is this slow growth, which makes the high degree of crystallinity of cellulose products like wood, comprehensible[2]. There is some evidence, that, as a rule, Nature builds up macromolecules with smaller deviations from the mean magnitude than synthetic polymers thus showing a steeper frequency-distribution curve.

The most striking example of this is to be found in the proteins, where even molecules of an exact magnitude are built, see Table 4, p. 30. In the fatty acids Nature prefers to build chains with an even number of C-atoms, especially C_{16} and C_{18}. In the case of xylene from straw, it was found[3] that 94% of the molecules had a polymerisation degree of 150. It must however be said that it is within the scope of laboratory technique to do the same, although until now our skill in this field is only limited. As an example, the preparation of polyethylene glycols $HO(CH_2CH_2O)_nH$ by American workers[4] may be pointed to, as they succeeded in obtaining a well defined chain length with n = 6 or 18 or 42 at will. The important consequences regarding the gels tructure of such substances is evident. In the absence of monomeric or other very small particles it would even be questionable whether one is dealing with gels in such cases.

The fact that in Nature certain polymers seem to have only a small variation in molecular size does not mean that the same also holds good for technical products, derived therefrom. In many[5] processes, necessary for modifying natural products, a serious breakdown of the macromolecules takes place and it may be asked which

[1] H. STAUDINGER und K. FEUERSTEIN, Ann., 526 (1936) 72.
[2] An extensive theory concerning the way in which this may occur is given by L. W. JANSSEN, Protoplasma, 33 (1939) 410.
[3] E. HUSEMANN, J. prakt. Chem., 155 (1940) 13.
[4] R. FORDYCE, E. L. LOVELL, H. HILBERT, J. Am. Chem. Soc., 61 (1939) 1905.
[5] Examples are known in which practically no breakdown occurs, even during a treatment such as the acetylation of cellulose, if this is carried out very mildly. Thus, STAUDINGER, Organische Kolloid-chemie, Braunschweig 1940, p. 115, finds for an amylum of polymerisation degree 560, a value of 540 after acetylation, and for a glycogen of polymerisation degree 5000, a value of 5300 after acetylation.

range of molecular sizes is obtained when a number of macromolecules is submitted to a degradation process. Leaving the mathematical treatment of this subject aside[1], we will confine ourselves to giving some results in Fig. 22 which have been obtained for the acetylation of cellulose. From 4000 g of the technical starting material, 850 g of the acetate with an average polymerisation degree of 300 was produced, the homogeneity of which was very reasonable, as shown at the right side of Fig. 22.
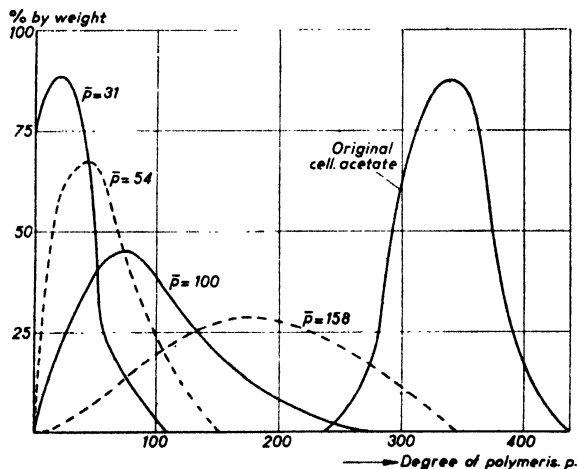
Fig. 22. Distribution of molecular sizes (mean degree of polymerisation $p = 31$, etc.) of specially prepared fractions of cellulose acetate

This material was again further depolymerised, under mild conditions with the aid of acetic acid, into four fractions with mean polymerisation degrees of 158, 100, 54 and 31 respectively. The distribution of the molecular sizes of these four fractions is shown at the left of Fig. 22, leading to the result that ultimately important quantities of very small particles are also formed. *Therefore, in technical products derived from natural polymers, the possibility of the isogel state must also be reckoned with, even when the starting material is free from low molecular products.*

---

[1] Cf., for a summary; H. MARK, *The General Chemistry of High Polymeric Substances*, Amsterdam 1940, p. 315.

III. THERMODYNAMICS OF LONG-CHAIN MOLECULES

J. J. HERMANS

Laboratory for inorganic and physical chemistry, University of Groningen.

§ 1. INTRODUCTION

a. Comparison between macromolecules and micromolecules

Systems containing macromolecules are, of course, subject to the ordinary thermodynamic laws. It is, in fact, quite important to realise that *there exists no essential difference of any kind between macromolecules and micromolecules.* All differences which do exist are of a quantitative, not a qualitative nature. We want to stress this point because it has met with antagonism at times, and it is still believed in certain circles, especially by those who have never come into close contact with colloid science, that this science is concerned with systems which are not, or not yet, accessible to a straightforward treatment on the basis of well-founded physico-chemical principles. There is no essential difference between a solution of micromolecules and one of macromolecules. Nor does there exist an essential difference between crystals composed of macromolecules and those consisting of micromolecules, and so on. It is only as a result of their particularly large size, that macromolecules sometimes exhibit a behaviour which does not have its counterpart in micro-physics, or that they occur in forms which do not exist in the micromolecular field.

Thus, for instance, in crystals consisting of micromolecules it never occurs that one molecule forms part of more than one crystal, a phenomenon which is by no means uncommon in solid macromolecular systems. To mention another item, the conditions which govern the association of macromolecules in solution are principally the same as those prevailing in solutions of micromolecules, but they may give rise to phenomena which do not occur in micromolecular systems in that they form coherent structures (network of scaffoldings) and thus lay the foundation to gel-formation. A further point of importance is the possibility of a large range of molecular sizes, which entails a great many phenomena which are unknown in micro-physics. Yet, all this, if interpreted rightly, emphasizes the similarity rather than the dissimilarity between macromolecules and micromolecules.

b. Scope of the present chapter

A number of macromolecules occur in the form of compact particles. This applies in particular to globular proteins. It is obvious, that the interaction of such a particle with its surroundings is confined to its outer surface, which means that the interaction energy is proportional to this surface. If the weight M of the particle is large, this means proportionality with $M^{2/3}$. A thermodynamical treatment of large

compact particles in effect works out to a thermodynamical treatment of interfaces and belongs to the first volume of this book rather than to the second. A few remarks concerning the thermodynamical aspects of the systems concerned can be found in Volume I. In the present chapter we shall confine the discussion to loosely built structures, whether randomly kinked or otherwise, where practically all parts of the molecule may interact to the same extent with surrounding molecules. A number of energy effects will then be proportional to the molecular weight (see below). There are intermediate cases, however, where the structure of the molecule is less simple. An example is given by strongly branched molecules such as starch or some polystyrenes, which form comparatively close-packed structures and are less easily accessible to a theoretical treatment. We shall mention these systems incidentally but will forgo detailed speculations concerning their behaviour.

As regards the subjects treated in this chapter, we do not aim at completeness. Colloid science has made use of thermodynamic principles in a variety of applications, and we can do no more than to select a few items which are of importance, trying to cast them into a coherent form as best we can. Further, our treatment will not be a purely thermodynamic one; molecular models will be discussed whereever this seems desirable.

§ 2.   THERMODYNAMIC FUNCTIONS. COEXISTING PHASES OF THE PURE SUBSTANCE

a. Symbols used

We will use the following symbols:

V, volume

S, entropy

U, internal energy

$F = U - TS$, free energy

$H = U + pV$, heat content or enthalpy

$G = U - TS + pV$, thermodynamic potential or GIBBS free energy.

Here T is the temperature and p the pressure. All quantities refer to one mole of the substance. In condensed systems (solids and liquids) the term pV may often be neglected compared with the other energy terms. We have then: $H \gtrsim U$; $G \gtrsim F$. In ideal gases $pV = RT$.

From the laws of thermodynamics it is possible to derive some relations between the quantities V, U, etc. We mention in particular.

$$V = \frac{\partial G}{\partial p}; \quad S = -\frac{\partial G}{\partial T}; \quad H = -T^2 \frac{\partial}{\partial T}\left(\frac{G}{T}\right) \tag{1}$$

Similar equations apply to *changes* in heat content, entropy, free energy, etc., and to the partial molar quantities in a mixture (compare p. 56).

For a detailed discussion of the thermodynamic functions mentioned and their applications the reader is referred to the textbooks on thermodynamics. We shall restrict ourselves to those points which are of importance to the subjects treated. The statistical aspect of the entropy is briefly discussed in section 6a p. 66.

We shall first examine the influence of the molecular weight on the values of U, S, etc. This subject was treated, among others, by HUGGINS[1]. We shall use the subscripts s and l for the solid and liquid state respectively.

---

[1] M. L. HUGGINS, *J. Phys. Chem.*, 43 (1939) 1083.

b. Thermodynamic functions with respect to degree of polymerisation

The heat content H of long-chain molecules in the solid state and in the liquid state alike are practically proportional to the degree of polymerisation P:

$$H_s \cong U_s = a_s + b_s P \qquad (2)$$

$$H_l \cong U_l = a_l + b_l P \qquad (3)$$

Here the coefficients a and b are functions of pressure and temperature but do not depend on P. The quantities a_s and a_l will depend on the terminal groups, while b_s and b_l are practically identical for all chains which are built up by the same monomeric groups. The equations (2) and (3) can be derived from considerations concerning the various energy contributions, vibrational, rotational and translational energy, bond energy and VAN DER WAALS energy, which all conform to equations of the type (2) and (3) in fair approximation[1]. The difference between b_s and b_l is chiefly due to a difference in rotational energy[2] and a small difference in the VAN DER WAALS energy.

From the heat content H we derive the heat capacity per mole at constant pressure:

$$c_p = \left(\frac{\partial H}{\partial T}\right)_p \qquad (4)$$

Using (2) or (3) we find for both solid and liquid an equation of the type

$$c_p = a + \beta P \qquad (5)$$

where a and β are independent of P. For the solid and the liquid hydrocarbons (normal paraffins) a plot of c_p against P is given in Fig. 1, showing a good agreement with form (5).
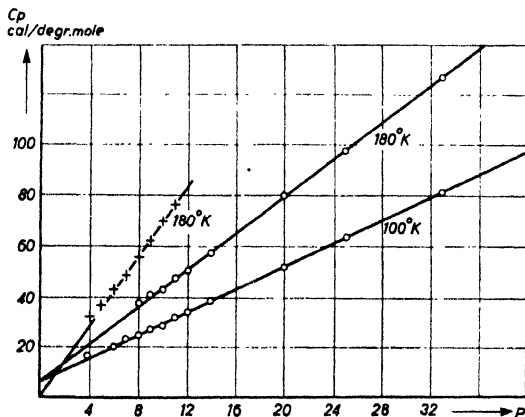
Fig. 1. Heat capacity[3] of crystalline hydrocarbons (o) at two temperatures and of liquid hydrocarbons (+).
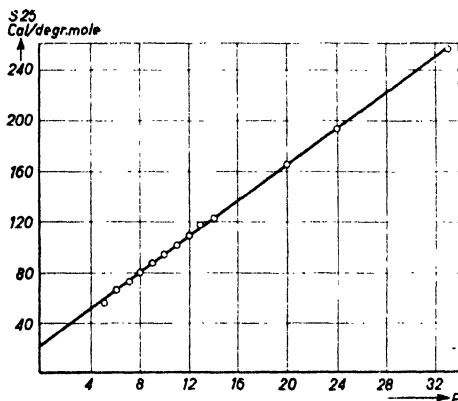
Fig. 2. Entropies[4] of liquid paraffins at 25° C.

[1] M. L. HUGGINS, *J. Phys. Chem.*, 43 (1939) 1083.
[2] W. E. GARNER and A. M. KING, *J. Chem. Soc. London*, (1936) 1368, 1372.
[3] G. S. PARKS, H. M. HUFFMAN and S. B. THOMAS, *J. Am. Chem. Soc.*, 52 (1930) 1032.
[4] H. M. HUFFMAN, G. S. PARKS and M. BARMORE, *J. Am. Chem. Soc.*, 53 (1931) 3876.

The entropy S of macromolecules in the liquid or solid state is represented with fairly good approximation by similar formulae[1]:

$$S_s = m_s + n_s P \tag{6}$$

$$S_l = m_l + n_l P \tag{7}$$

Here m and n do not depend on the degree of polymerisation P. A plot of S_l against P for some lower members of the normal paraffins is given in Fig. 2. The expressions for the free energy F and the GIBBS potential G can be obtained from H and S, using the definitions in section 2a.

It is obvious that the results obtained so far may be summarized by saying that all thermodynamic functions *per mole* of the pure substance are linear with respect to P, which means that the thermodynamic functions *per gram* at high values of P become independent of the degree of polymerization.

c. The melting process; melting point and heat of melting

c. 1 *Melting point*

If two phases of a substance are in equilibrium with each other, the GIBBS potential in the one phase equals that in the other. For instance, in the equilibrium between liquid and solid

$$G_s(p, T) = G_l(p, T) \tag{8}$$

This equation determines the transition temperature T_{tr} at a given pressure. Introducing the symbol $\varDelta$ for the change in a quantity when one phase is converted to the other, we have $\varDelta G = 0$. Using the definition $G = H - TS$ this gives:

$$T_{tr} = \frac{\varDelta H}{\varDelta S} \tag{9}$$

In the equilibrium between solid and liquid state T_{tr} is the melting point T_f and $H = H_l - H_s$ represents the heat of melting per mole. From the equations (2) and (3) it follows that this heat of melting is a linear function of P, as is borne out by experiment[2]. We may also express this by saying that the heat of melting per gram is independent of P if P is large. It is to be noted in this connection that the coefficients a and b in equations (2) and (3) are functions of the temperature. The experimental value of $\varDelta H$, however, refers to the melting point of the particular substance concerned. To check the relation between $\varDelta H$ and P one should, therefore, first reduce all $\varDelta H$- and $\varDelta S$-values to the same temperature, which can be done with the aid of the experimental values of c_p. In fact,

---

[1] On p. 73 it is shown that the entropy of liquid or solid amorphous polymers contains a term $R \ln P$ but the value of $\ln P$ changes but slowly with P and may be considered as practically constant over a considerable range of P-values.

[2] W. E. GARNER, F. C. MADDEN and J. E. RUSHBROOKE, *J. Chem. Soc.* (1926), 2491; A. M. KING and W. E. GARNER, *J. Chem. Soc.* (1936) 1368, 1372. In the paraffin derivatives examined by these authors, the derivatives with an even number of C-atoms must be distinguished from those with an odd number of C-atoms.

$$H(T) = H(T') - \int_T^{T'} c_p \, dT$$

for either the solid or the liquid state. At high values of the degree of polymerization no such reduction is required, since the melting point soon becomes independent of P. In fact, using the equations 2—7 in (9) we get

$$T_f = \frac{a' + b'P}{m' + n'P} \tag{10},$$

where $a'b'm'$ and n' are independent of P. An equation of this form (with values of $a'b'm'$ and n' independent of temperature) was used by GARNER[1] and co-workers to represent the melting point of a number of paraffin derivatives. If P is sufficiently large, T_f becomes b'/n'. Since both b' and n' are independent of the terminal group, this limiting melting point at high values of P is also independent of the terminal groups[1].

c. 2. Melting range

Does it follow from this result that a mixture of polymer homologues has a melting point which is independent of the composition? This is a question of certain interest, which has received much attention in past and recent years. Most macromolecular substances consist of a series of polymer homologues. In other words, they do not represent a one-component system at all. The fact that the various components in this mixture have the same melting point does not in itself guarantee that the melting point of the mixture is the same as that of the separate components. As a matter of fact, we shall show later (p. 75) that one must expect on purely thermodynamic grounds that the melting point of a polymer homologous mixture may be somewhat different from that of the pure components, and that, consequently, the mixture will show a certain melting interval: the melting point is not sharp. This behaviour depends essentially on the entropy of mixing (see p. 66 and 75). In practice, however, it is almost impossible from the experimental temperature interval to draw any quantitative conclusions of theoretical importance because there exist several other causes which may make the melting point not sharp. One of these will become obvious when we consider the sharp melting point of micromolecular substances.

In the crystalline state the atoms are oscillating in a potential trough at fixed positions in the lattice. With increasing temperature the amplitude of these oscillations increases until a certain fraction of the atoms have acquired a sufficient amount of oscillatory energy to break out of their stable position, thus disturbing the lattice. As a result of this disturbance the potential trough of the atoms in the immediate neighbourhood is flattened and these atoms also are thus enabled to take part in the melting process. This is often expressed by saying that the melting process is a *cooperation phenomenon*: each atom which takes part in it increases the mobility of

---

[1] W. E. GARNER, F. C. MADDEN and J. E. RUSHBROOKE, *J. Chem. Soc.* (1926) 2491; A. M. KING and W. E. GARNER, *J. Chem. Soc.* (1936) 1368, 1372. In the paraffin derivatives examined by these authors, the derivatives with an even number of C-atoms must be distinguished from those with an odd number of C-atoms.

the surrounding atoms, and consequently the whole process can play itself out within a very narrow temperature range. In macromolecular substances, however, this cooperation is less effective, since the molecules extend over much larger distances. If a small region in the lattice is disturbed, the molecule may still be fixed at other points, and the mobility of the "molten part" of the molecule is thereby reduced.

Still more effective is the fact that the lattice in a crystal consisting of macromolecules often shows a less complete long range order[1] than that of micromolecules. This is quite intelligible if one realizes the great difficulties which must be overcome if these long molecules are to arrange themselves in a regular lattice. Accordingly, the potential barriers which limit the oscillations in the lattice are not all of equal height, which means that the energies needed to break out of the potential holes comprise a continuous range, with a corresponding range of melting points. These views are, of course, closely related to the suggestion which was put forward by DOSTAL[2], according to which the *size* of the crystals influences the melting point. In fact, if the crystals are very small, the crystalline order extends over a small number of atoms, with the result that the lattice energy is different from that in the larger crystals.

Finally, we refer to section 2d, in which we discuss a further reason why the melting interval is comparatively large.

d. Crystallization in macromolecular solids

So far we have tacitly assumed the existence of two well-defined phases. In many solid macromolecular substances, however, the situation is much less simple.

Consider the equilibrium between ice and water. This is a two-phase system and, accordingly, has but one degree of freedom: if at constant pressure the temperature is changed, either all ice is melted or all water is frozen. Now, X-ray analysis has shown that many macromolecular solids contain crystals, although at the same time a considerable part of these substances is in the amorphous state (compare Chapter XII, p. 495). The fundamental question arises, whether such a system must be considered as a one-phase or a two-phase system. If it were an equilibrium two-phase system, it could (at a given pressure) only exist at one temperature. We are either to conclude, therefore, that there is but one phase, or that there are two phases which are not in equilibrium with each other. The structure which is fairly generally accepted nowadays is one in which crystalline regions alternate with amorphous parts of the material. One molecule may form part of one or more crystals and of the interjacent amorphous regions as well. For this reason the system may show the properties of a one-phase system, in which the crystals represent regions with a pronounced order, in contrast with the amorphous parts where this order is less complete. There would then be no essential difference between the structure of such a solid and that of ordinary liquids, the only difference being that the crystalline order in the crystalline regions of the solid is more pronounced and extends over greater distances than that in the well-ordered parts of a liquid.

---

[1] Studies on the heat content of crystals with disturbed lattice have been made, among others, by R. FRICKE, *Z. Elektrochem.*, 44 (1938) 291; 46 (1940) 90, 491, 641; *Z. Physik. Chem.*, A 183 (1938) 165, 177; B. 37 (1937) 60; 39 (1938) 476.
[2] H. DOSTAL, *Osterr. Chem. Ztg.*, 41 (1938) 20.

A number of compounds such as caoutchouc and related substances are amorphous at room-temperature but give rise to X-ray interferences if they are stretched. These interferences appear at 100—300% stretch; their intensity increases with increasing stretch, showing that the percentage crystalline matter increases. The phenomenon is completely reversible and therefore cannot be described by the existence of two separate phases. For, in a two-phase equilibrium there would be but one temperature at a given stretch where the two phases are coexistent (melting point of the substance). In the present case, it is true, we may also introduce a so-called melting point, which changes with the degree of stretch. But this melting point is a more or less *artificial* one: it represents the temperature at which the crystal interferences in the X-ray diagram practically disappear. The solid is not completely crystalline below this melting point as it should be if it behaved as a two-phase equilibrium system.

It was shown by FRITH and TUCKETT[1], that the one-phase model may explain the comparatively long melting range (compare, however, p. 53). The argument given amounts to the following. If the polymer consists of two separate phases, i.e., if a molecule belongs either to the crystalline or to the amorphous part, both the entropy and the heat content of the system will be *linear* functions of the fraction (Θ) of crystalline substance:

$$S = (\Theta)S_{cr} + (1 - \Theta)S_{am}; \quad H = (\Theta)H_{cr} + (1 - \Theta)H_{am} \tag{11}$$

The melting point T_f is determined by the condition of minimum free energy:

$$\frac{\partial H}{\partial (\Theta)} - T \frac{\partial S}{\partial (\Theta)} = 0 \tag{12}$$

In the two-phase system this gives obviously

$$T_f = \frac{H_{am} - H_{cr}}{S_{am} - S_{cr}}, \tag{13}$$

which means that T_f is independent of (Θ). If, however, a single molecule may extend over both crystalline and amorphous parts, the average number of atoms in an amorphous chain changes with the fraction (Θ). The entropy S_{am} of the amorphous material, however, depends on the number of atoms in the chain (compare p. 73). In that case the equilibrium condition (12) assumes a less simple form:

$$T_f (S_{am} - S_{cr}) = H_{am} - H_{cr} + (1 - \Theta) T_f \frac{\partial S_{am}}{\partial (\Theta)} \tag{14}$$

Thus the melting point will no longer be independent of Θ, and will show a continuous shift while the substance is melting. If for the entropy of the amorphous chains one uses FLORY's expression, which will be discussed in section 6d p. 73, one finds a melting range which actually has the right order of magnitude (FRITH and TUCKETT, *loc. cit.*).

However, it is to be noted, that the one-phase concept need not apply to all solid polymer substances. In cellulose no measurable influence of stretch on crystallinity is observed[2], let alone a reversible one. It is quite possible that the attainment

[1] E. M. FRITH and R. F. TUCKETT, *Trans. Faraday Soc.*, 40 (1944) 251.
[2] O. KRATKY and A. SEKORA, *Kolloid-Z.*, 108 (1944) 169; see also the chapter on gels, p. 613.

of a crystallisation equilibrium in cellulose derivatives is practically infinitely slow. For that matter, very low crystallisation velocities are often observed in caoutchouc and similar substances[1], where, moreover, the melting point often changes with the preliminary treatment of the sample.

§ 3. GENERAL PROPERTIES OF SOLUTIONS

Since the solutions of macromolecular substances are of particular importance to the studies of their properties and structures, a major part of this chapter will be devoted to solutions in general and those of macromolecules in particular.

So far, we have only considered the pure polymer substance. Strictly speaking, such a substance contains more than one component, because it always contains molecules of various molecular weights. This does not, however, affect the heat content H, since per gram of substance this heat content is independent of the size of the molecules, provided the mol. weight is high. The possible role of the *entropy* in the properties of a polymer homologous mixture will be briefly considered in section 6e p. 75.

We shall now become concerned with mixtures containing various types of molecules (not only various sizes), and we shall first give a short survey of some general thermodynamic properties.

a. Partial molar quantities; free energy and heat of dilution (and solution)

a. 1. Considering a mixture with heat content H, GIBBS free energy G, etc., containing $N_0 \, N_1 \, N_2 \ldots \ldots$ moles ot species $0,1,2 \ldots$ we define as partial (molar) heat content, molar free energy[2], etc. of these species,

$$h_i = \frac{\partial H}{\partial N_i}; \quad g_i = \frac{\partial G}{\partial N_i}; \quad s_i = \frac{\partial S}{\partial N_i}; \quad v_i = \frac{\partial V}{\partial N_i} \tag{15}$$

The physical meaning for example, of the molar volume v_i is the increase in the volume of a large amount of mixture when 1 mole of the pure component i is added to it. From the relations (1) on p. 50 it follows that these molar quantities conform to analogous relations:

$$h_i = g_i + T s_i \; ; \quad h_i = \frac{\partial g_i/T}{\partial 1/T}; \quad s_i = -\frac{\partial g_i}{\partial T} \; ; \quad v_i = \frac{\partial g_i}{\partial p} \tag{16}$$

The functions G, H, S, V are directly proportional to the total amount of the mixture. If all numbers N_i are doubled, these functions also become twice as large. It follows that in a mixture of two components, o and r, the change of g_o with concentration x is connected with the change in g_r according to the well-known GIBBS-DUHEM (often also called DUHEM-MARGULES) relation[3]:

[1] A. VAN ROSSUM and J. LOTICHIUS, *Kautschuk*, 5 (1929) 2.
[2] It can be shown that $(\partial G/\partial n)_{p,T} = (\partial F/\partial n)_{v,T}$, i.e., the molar GIBBS free energy is equal to the molar free energy. No confusion can arise, therefore, if we omit "GIBBS" for brevity.
[3] In the following we will often use the subscript o to denote the solvent and the subscript r to denote the (polymer) solute. Here r may be considered as abbreviation of "rubber" (G. GEE, *Trans. Faraday Soc.*, 38 (1942) 276; 40 (1944) 463; *Trans. I. R. I.*, 18 no 6).

$$(1 - x)dg_o + x \, dg_r = 0 \tag{17}$$

where x is the *molar ratio*:

$$x = \frac{N_r}{N_o + N_r} \tag{18}$$

The quantities h_o and h_r are related by an equation which is similar to (17); so are s_o and s_r, etc.

a. 2. Quantities of great practical interest are the free energy and heat of dilution, and the free energy and heat of solution. The heat of *dilution* is the increase in the heat content of the whole system when 1 mole of solvent is transferred from a certain mass of pure solvent to an infinite amount of the mixture. In other words, when h_o° represents the heat content per mole of pure solvent and h_o the (partial) molar heat content of the solvent in the mixture, the heat of dilution is

$$\mathit{l}h_o = h_o - h_o^{\circ} \tag{19}$$

It follows from this definition that a positive heat of dilution means that heat is absorbed from the surroundings when diluting the mixture. Similarly, the GIBBS free energy and the entropy of dilution are given by the equations

$$\mathit{l}g_o = g_o - g_o^{\circ} \; ; \quad \mathit{l}s_o = s_o - s_o^{\circ} \tag{20}$$

It will further be obvious that the relations (16) also apply to the difference between h_o and h_o°, etc. In other words, the heat of dilution can be derived from the GIBBS free energy of dilution and its temperature coefficient, while the entropy $\mathit{l}s_o$ of dilution follows from $\mathit{l}g_o$ and $\mathit{l}h_o$ or from the temperature coefficient of $\mathit{l}g_o$.

Finally, we mention the free energy and heat of *solution*, which are completely analogous to those of dilution. For instance, the GIBBS free energy of solution is the increase in GIBBS free energy when transferring 1 mole of pure solute from a mass of pure solute to a large quantity of the mixture. In view of the relation (17) between g_o and g_r, the free energy of solution can be derived from that of dilution. Integrating (17) we get

$$g_r - g_r^{\circ} = - \int_1^x dx \, \frac{1-x}{x} \, \frac{dg_o}{dx}, \text{ or } \quad \mathit{l}g_r = - \int_1^x dx \, \frac{1-x}{x} \, \frac{d \, \mathit{l}g_o}{dx} \tag{21}$$

with similar equations relating $\mathit{l}h_r$ to $\mathit{l}h_o$, and $\mathit{l}s_r$ to $\mathit{l}s_o$. Conversely, the value of $\mathit{l}g_o$ may be expressed in that of $\mathit{l}g_r$:

$$\mathit{l}g_o = - \int_0^x dx \, \frac{x}{1-x} \, \frac{d \, \mathit{l}g_r}{dx} \tag{22}$$

In practice the integrations will usually have to be carried out graphically [1]. The quantities $\mathit{l}h_o$, $\mathit{l}h_r$ etc., which have been introduced here, are directly accessible to experiment, and have often been used as a means to study the thermodynamic properties of mixtures and as a check on theoretical formulae. We will give a short account of the experimental methods in the following subparagraph.

[1] See, for instance, G. GEE and L. R. G. TRELOAR, *Trans. Faraday Soc.*, 38 (1942) 147.

b. Determination of thermodynamic quantities in a mixture

b. 1. *Vapour pressure*

An experimental study of thermodynamic quantities in solutions of macro-molecules has almost always been based on either the vapour pressure or the osmotic pressure [1]. Let us first consider the vapour pressure method. If the solution is in equilibrium with the vapour of the pure solvent, the molar free energy of the solvent must be the same in the two phases. Denoting the quantities pertaining to the solvent by a subscript o and the vapour phase by a stroke,

$$g_o = g'_o$$

If the vapour has the properties of an ideal gas and if p is the pressure, we have $g'_o = f + RT \ln p$, where f is a function of temperature alone. Consequently the molar free energy in the solution is

$$g_o = f + RT \ln p$$

Considering, in addition, the equilibrium between the pure solvent and its vapour at the same temperature and at pressure p^o, and writing (as usual) g_o^o for the molar free energy of the pure solvent, we must have

$$g_o^o = f + RT \ln p^o$$

Thus, from the vapour pressure p of the mixture and the vapour pressure p^o of the pure solvent we derive the GIBBS free energy of dilution

$$\Delta g_o = g_o - g_o^o = \Delta h_o - T \Delta s_o = RT \ln \frac{p}{p_o} \qquad (23)$$

At this junction let us anticipate a result, which will be obtained in a later section (p. 67), bearing upon solutions in which the entropy of dilution is given by the simple expression:

$$\Delta s_o = - R \ln(1 - x) \qquad (24)$$

where x is the mole fraction. The vapour pressure will then be determined by the equation

$$\frac{p}{p_o} = (1 - x) \, e^{\Delta h_o / RT} \qquad (25)$$

With zero heat of dilution this represents RAOULT's law (straight line in Fig. 3), showing that RAOULT's law results entirely from entropy effects. The upper curve in Fig. 3 represents a case where $\Delta h_o > 0$ (heat is absorbed on mixing), while the lower curve shows $\Delta h_o < 0$. If Δs_o is not ideal, it is obvious from equation (23) that an entropy of dilution which is higher than the ideal one results in a vapour pressure below the straight line, and vice versa. It will become apparent in section d. 2, that the deviations from RAOULT's law in solutions of long-chain molecules are due first and foremost to deviations of the entropy of mixing from the ideal law.

---

[1] The recent studies on light scattering in polymer solutions lead to principally the same results as the osmotic pressure method. Compare the chapter on molecular weight determinations, p. 146.

However great the deviations from RAOULT's law may be, in the limit of zero concentration it will always hold good. This is why the slope of the curves in fig. 3 approaches the value 1 at low values of x. This is due to the fact that the heat of dilution Δh_o becomes a higher order of zero than x itself when x approaches zero (compare p. 66 and 70).

b. 2. Osmotic pressure

A second way of deriving Δg_o is from the determination of osmotic pressure π. The osmotic equilibrium between the solution and the solvent requires that the molar free energy of the solvent in the mixture is equal to that in the pure solvent. Taking into account that the pressure in the mixture is $p + \pi$ if that in the pure solvent is p, we get

$$g_o\,(p + \pi) = g_o^o(p)$$

Fig. 3.   Vapour pressure plotted against mole fraction x.

Using the relation $\partial g_o / \partial p = v_o =$ molar volume of solvent, and assuming that v_o is practically independent of the pressure, this gives

$$g_o + \pi v_o = g_o^o \text{ , or} \qquad \Delta g_o = \Delta h_o - T\,\Delta s_o = - \pi v_o \tag{26}$$

Here again, let us consider the special case of "ideal" solutions (see preceding sub-paragraph), assuming that the heat of dilution is zero and the entropy of dilution conforms to equation (24). This gives

$$\pi = - \frac{RT}{v_o}\ln\,(1 - x) \tag{27}$$

In solutions of macromolecules the mole concentration x will practically always be very small, since the molecular weight M of the solute is high. Introducing the concentration c (grams in unit volume) the equation (27) will then take the form of VAN 'T HOFF's law

$$\pi = \frac{RT}{v_o}x \cong \frac{RT}{M}c \tag{28}$$

This shows that in solutions of macromolecules all deviations from VAN 'T HOFF's law must be attributed to the non-ideal character of the solution. We will show on p. 74, that the entropy of dilution Δs_o in solutions of long-chain molecules is considerably greater than the ideal one. According to equation (26) this will lead to an osmotic pressure which is higher than VAN 'T HOFF's value.

b. 3. The heat of dilution Δh_o may be measured directly in the calorimeter[1], or may be determined by the difference between the heats of vaporisation from the solution and from the pure solvent. An indirect method consists of using the relation (16):

$$\Delta h_o = \left| \frac{\partial\,(\Delta g_o\,/T)}{\partial\,1/T} \right|_{p.x}$$

This involves the determination of Δg_o (i. e., vapour pressure or osmotic pressure) at various temperatures. Finally the *entropy* of dilution may be obtained from $\Delta g°$ and Δh_o (since $\Delta g_o = \Delta h_o - T \Delta s_o$) or also from the temperature coefficient of Δg_o :

$$\Delta s_o = - \left| \frac{\partial\,\Delta g_o}{\partial T} \right|_{p.x}$$

§ 4. SIMPLE MODEL OF POLYMER SOLUTIONS

All relations obtained so far are of a general nature (except for the equations 24, 25, 27, 28 which apply to ideal solutions only). Equations relating the thermodynamic quantities of a solution to the concentrations of the components, the temperature, etc. can only be obtained by considering special models. The model of a liquid which is most appropriate to the purpose is that which assumes a system which is almost close-packed with a rather pronounced short-range order[2]. This means that within a volume of molecular dimensions the arrangement is almost regular and differs but slightly from that in a crystal lattice. Each atom is surrounded by a certain number of closest neighbours; this number is the *coordination number* of the lattice. This short-range order is due essentially to the close packing in the liquid: the position of the molecules surrounding a given atom is limited, on the one hand by the size of this atom, and on the other hand by the attractive forces.

As a result of small deviations from the regular order, i.e., small derangements of the "lattice", there exists no long-range order in the liquid: the correlation between the positions of two molecules becomes progressively less pronounced as the distance between these molecules increases.

In section 6d, p. 72, we shall be concerned with the entropy of polymer solutions. This entropy is derived from the number of possible molecular arrangements in the solution. It is here, where the simple model of a more or less regular lattice has been of such great value. Applying this model to polymer solutions, it is assumed that each polymer molecule may be treated as if it consists of a large number n of *segments*. These segments are approximately equal in size and shape to a single molecule of the solvent and may replace the solvent molecules in the "lattice". This does not necessarily imply that the polymer molecules are of the randomly kinked type (compare chapter on random coiling, p. 93). The treatment applies, in princip'e, to polymer molecules of any shape. But it is obvious, that the conditions underlying this treatment are more easily met with by randomly coiled polymer molecules than,

---

[1] An interesting compensation calorimeter was used by E. CALVET, *J. chim. phys.*, **35** (1938) 69; *Compt. Rend.*, **212** (1941) 542.

[2] R. H. FOWLER and E. A. GUGGENHEIM, *Statistical thermodynamics*, p. 319, Cambridge 1939.

for instance, by rigid rod-shaped particles. For, with rigid particles it depends on the lattice type of the solvent whether it will be at all possible simply to replace a certain number of solvent molecules, on lattice sites, by a solute molecule. It is, in particular, the assumption that the lattice is not seriously disturbed, which characterises the model, and which to a certain extent determines its somewhat artificial nature. It will become apparent in section 6d, p. 72, however, that a more general treatment would meet with considerable difficulties.

The model implies further, that the volume is not materially changed on mixing. If there are N_o solvent molecules and N_r polymer molecules, each containing n segments, the volume fractions are

$$q_o = \frac{N_o}{N_o + nN_r} ; \quad q_r = \frac{nN_r}{N_o + nN_r} \quad (29)$$

We shall make use of these relations in later sections. The number n is given experimentally by the ratio between the molecular volume of the solute and that of the solvent. If the molecular weight of the polymer is known, this ratio is directly accessible to experiment.
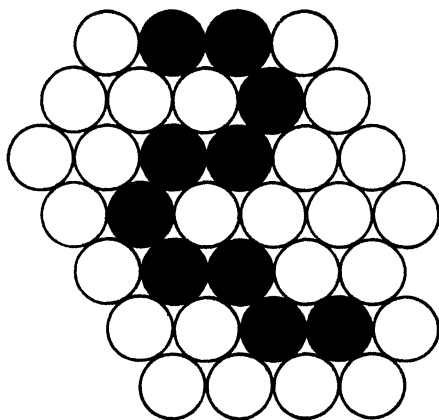
Fig. 4. Segments of polymer molecule on lattice sites in liquid.

§ 5. HEAT OF MIXING

Since, evidently, the heat content H of a mixture will depend on the interaction between the various components, there can be no question of a completely general treatment. All we can do is to consider a few specially simple limiting cases. The simplest case is, of course, that in which the components mix without any change in the heat content. This condition is approximately met with in a few cases, when the molecules are sufficiently alike.

Following the usual nomenclature, we call heat of mixing $\triangle H$ the difference between the heat content of the mixture and that of the separate components. Thus, a positive heat of mixing means that heat is absorbed on mixing. Until recently, it was fairly generally believed that polymer molecules could be dissolved in a liquid only if the heat evolved on mixing was zero or positive [1]. The reasons for this belief will become clear later (see p. 71). Recent investigations have shown, however, that a great many macromolecules dissolve in liquids of low molecular weight with absorption of heat. This can be explained by the considerable entropy of mixing (see p. 73).

Interpreting the heat of mixing on a molecular basis, a large negative value of $\triangle H$ clearly means that the molecules of the solute are strongly attracted by the solvent molecules. Conversely, if $\triangle H$ is positive, energy is needed to achieve mixing, which

[1] G. V. SCHULZ, Z. physik. Chem., A 180 (1937) 1.

means that solvent-solvent contacts and solute-solute contacts will be established in preference to contacts between solvent and solute molecules. The tendency of the system to reach its state of lowest energy will then inhibit a completely random distribution. A simple theoretical treatment of this effect does not exist (see the references on page 76), and we will here only consider those cases in which $\triangle H$ is comparatively small, assuming that it does not materially affect the randomness of molecular distribution. On this assumption we can give a simple formula for the heat of mixing, if we assume further, that we need only consider the interaction energy of nearest neighbours.

a. Derivation of formula

We adopt the simple model of the preceding section. The following considerations could easily be put on a somewhat wider basis, if required. However, we shall anyhow make use of this model in the calculation of the entropy of mixing (p. 72), and it is not our aim here to investigate the minimum conditions required for the following relations to be applicable.

If there are N_o solvent molecules and N_r solute molecules, each consisting of n segments, there are $N_o + nN_r$ sites, of which N_o are occupied by solvent molecules and nN_r by polymer segments. Let z be the coordination number. If mixing may be considered as random, each solvent molecule is surrounded on the average by $zN_o(N_o + nN_r)^{-1}$ solvent molecules and by $znN_r(N_o + nN_r)^{-1}$ polymer segments. If $-a_{oo}$ represents the VAN DER WAALS energy of two neighbouring solvent molecules and $-a_{or}$ that between a solvent molecule and a polymer segment, then, since all interaction except that between nearest neighbours is neglected, the contribution of the solvent molecules to the interaction energy becomes

$$U_o = - N_o [a_{oo}zN_o + a_{or}znN_r] (N_o + nN_r)^{-1}$$

Similarly each polymer segment is surrounded by z adjacent sites, of which $znN_r(N_o + nN_r)^{-1}$ are occupied by polymer segments and $zN_o(N_o + nN_r)^{-1}$ by solvent molecules [1]. If $-a_{rr}$ is the VAN DER WAALS energy of neighbouring segments, we find the following contribution of the polymer molecules to the interaction energy:

$$U_r = - nN_r [a_{or}zN_o + a_{rr}znN_r] (N_o + nN_r)^{-1}.$$

Since each interaction energy is in this way counted twice, the energy $U_o + U_r$ must be halved. Comparing with the energy $-\tfrac{1}{2}za_{oo}N_o - \tfrac{1}{2}za_{rr}nN_r$, we obtain the energy difference

$$\triangle H = \tfrac{1}{2}z [a_{oo} + a_{rr} - 2a_{or}] \frac{N_o\, nN_r}{N_o + nN_r}$$

If we abbreviate

$$\beta = \tfrac{1}{2}z [a_{oo} + a_{rr} - 2a_{or}], \tag{30}$$

---

[1] In this simple form this is only true if the chain molecules are of such a nature that successive segments in a chain do not occupy adjacent lattice sites. In most cases this condition is violated and the calculation becomes less simple, without, however, seriously affecting the final result (32), provided the coordination number z is not too small.

the heat of mixing assumes the simple form

$$\triangle H = \beta \frac{N_o \, nN_r}{N_o + nN_r}.$$ (31)

If v_o is the molecular volume of the solvent, and thus nv_o that of the polymer, the expression (31) represents the heat of mixing for a volume $v_o (N_o + nN_r)$ of the mixture. Using the equations (29), this means that the heat of mixing per unit of volume is

$$(\triangle H)_1 = \frac{\beta}{v_o} \varphi_o \, \varphi_r$$ (32)

This is the simple parabola shown in Fig. 5. An example is given in the heat of mixing in the system carbon disulfide — dihydromyrcene (Fig. 6), which shows good agreement with the theoretical formula. Similar results were obtained in mixtures of dihydromyrcene with benzene, toluene, heptane and carbon tetrachloride[1].

b. Discussion

The coefficients a_{oo} and a_{rr} are related to the latent heat of evaporation at constant volume of the pure components. For, obviously, $\frac{1}{2}z \, a_{oo} N_o$ represents the total cohesive energy of N_o solvent molecules, and the ratio a_{oo}/v_o is a measure of the *cohesive energy density* of the pure solvent. It was suggested by HILDEBRAND[2] that a_{or} would be related to a_{oo} and a_{rr} as follows:

$$a_{or} = \sqrt{a_{oo} \, a_{rr}}$$ (33)
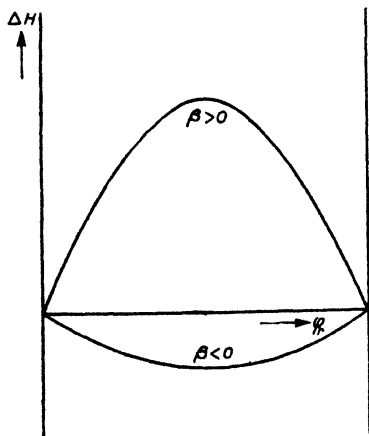
Fig. 5. Heat of mixing per unit of volume according to equation (32).
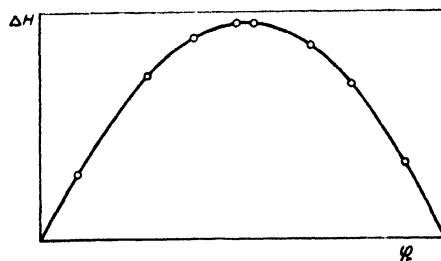
Fig. 6. Heat of mixing of CS_2 and dihydromyrcene[1].

Using equation (30), this leads to the following simple equation for β:

$$\beta = \frac{1}{2}z \, (\sqrt{a_{oo}} - \sqrt{a_{rr}})^2$$ (34)

[1] J. FERRY, G. GEE and L. R. G. TRELOAR, *Trans. Faraday Soc.*, 41 (1945) 340.
[2] J. H. HILDEBRAND, *Solubility*, p. 73, Reinhold 1936; J. H. HILDEBRAND and P. WOOD, *J. chem. Physics*, 1 (1933) 817.

where a_{oo} and a_{rr} are proportional to the cohesive energy densities of the pure components. In this case the heat of mixing is essentially positive. Negative values of $\triangle H$ cannot be explained on HILDEBRAND's assumption, but this does not necessarily exclude the applicability of the above equations with negative values of β.

If the expression (31) is differentiated with respect to N_o we obtain the heat of dilution

$$\triangle h_o = \partial \triangle H / \partial N_o = \beta \varphi_r{}^2 \tag{35}$$

Following HUGGINS, FLORY and others, we shall denote this expression as VAN LAAR heat of dilution. An equation of this type has been used by a number of workers. In those cases where $\triangle h_o$ is positive and not too large, it represents a useful overall formula which accounts fairly well for the dependence on concentration.

In other cases, where $\triangle h_o$ is negative and large, no such simple formula holds good. Now, large negative values of the heat of mixing often represent an indication for specific interaction: the solvent is strongly bound by the polymer in rather well-defined compounds. For example, cellulose nitrate binds 6 molecules of acetone per glucose unit (see p. 91); cellulose gives hydrates with water (see p. 539), etc. Such pronounced specific interactions cannot be described by a simple theory which assumes random mixing. We should like to suggest, however, that the formulae obtained here might even be used in these special cases, provided they are applied to the polymer-solvent compound rather than to the pure polymer itself. In other words, one would have to distinguish two stages in the dissolution process. In the first stage the specific attraction is predominating, leading to a rather well-defined compound; the second stage may be considered as a dissolution of this compound in the pure solvent. As a matter of fact, the heat of dilution in the system cellulose nitrate-acetone is negative at low concentrations of acetone but assumes small positive values at higher concentrations (compare p. 91).

It is assumed in the above formulae that the interaction (a_{oo}, a_{or}, a_{rr}) is not changed on changing the composition. This is one of the reasons why the calculation is confined to action between nearest neighbours. In general the interaction will depend on factors such as the dielectric constant of the mixture. For this reason the formula (35) must be considered as a first approximation. As pointed out by GEE and TRELOAR[1], deviations from this simple law may sometimes be explained on the basis of LANGMUIR's theory. The assumption that only closest neighbours contribute to the interaction energy was further specified by LANGMUIR[2], by saying that the decisive factor is the molecular surface rather than the volume. If a_o and a_r are the molecular surface areas per unit of volume of the two components, the molecular surface fractions are in the ratio

$$\frac{A_o}{A_r} = \frac{a_o \varphi_o}{a_r \varphi_r} ; \quad A_o + A_r = 1 \tag{36}$$

The heat of dilution becomes $\beta' A_r{}^2$ instead of (35). This gives

$$\triangle h_o = \beta' \left[\frac{\varphi_r}{\dfrac{a_o}{a_r} + (1 - \dfrac{a_o}{a_r}) \, \varphi_r} \right]^2,$$

which shows that instead of (35) we may expect a relation of the type

$$\triangle h_o = \frac{\beta \varphi_r{}^2}{(1 + \gamma_o \, \varphi_r)^2} \tag{37}$$

[1] G. GEE and L. R. G. TRELOAR, *Trans. Faraday Soc.*, 38 (1942) 147.
[2] I. LANGMUIR, *Colloid Symposium Monograph*, 3 (1925) 48.

where γ_o (either positive or negative) will in practice be a purely empirical constant. An example is given in Table I, showing the heat of dilution in the system rubber-benzene[1]. Obviously $\triangle h_o \, \varphi_r^{-2}$ is only constant over a very limited concentration range. The product $\triangle h_o \, \varphi_r^2 \, (1 - 0.60 \, \varphi_r)^2$, however, remains constant throughout the entire concentration range (see also Fig. 7). In Table 1 this product is denoted by C for shortness.

TABLE 1

HEAT OF DILUTION (CAL PER G BENZENE) IN RUBBER-BENZENE AT $25^\circ C$[2]

φ_r	$= 10^{-6}$	10^{-4}	10^{-2}	0.05	0.12	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
$\varLambda h_o$	$= 7 \cdot 10^{-13}$	$7 \cdot 10^{-9}$	$7 \cdot 10^{-5}$	$1 \cdot 9 \cdot 10^{-3}$	$1 \cdot 2 \cdot 10^{-2}$	$9 \cdot 5 \cdot 10^{-2}$	0.19_5	0.36	0.63	1.0_4	1.7	2.7	4.5
$\varLambda h_o \, \varphi_r^{-2}$	$= 0.7$	0.7	0.7	0.76	0.83	1.06	1.22	1.44	1.75	2.1	2.7	3.3	4.5
C	$= 0.7$	0.7	0.7	0.72	0.72	0.71	0.71	0.71	0.72	0.71	0.71	0.71	0.72

c. Heat of dilution at low concentrations

It will be obvious that equations (35) and (37) are not necessarily restricted to the special model used. In fact, equations of this type will always be obtained when mixing is random and interaction is confined to nearest neighbours[3]. With small values of the concentration $\triangle h_o$ becomes proportional to φ_r^2, i. e., *a small quantity of a higher order than the concentration itself,* a result which evidently is not affected by corrections of the LANGMUIR type (equation 37). In section 6 b. 3, p. 70 this result will be derived from a general thermodynamic consideration. It is of such great importance in its application to the properties of dilute solutions, that we shall also indicate its molecular origin.

Fig. 7. $\varLambda h_o$ from Table 1

plotted against φ_r^2 and against $\varphi_r^2 \, (1 - 0.6 \, \varphi_r)^{-2}$

If we transfer a molecule from the pure solvent (b in Fig. 8) to the solution (a), we will gain a heat $\triangle h_o$ = heat of dilution. If a solvent molecule from (b) is placed in a position whose vicinity contains only solvent molecules, the heat gained in the process is, of course, zero. But also, if the solvent molecule from (b) is brought into a position where it can interact with a single solute molecule only, the heat content of the whole system remains unaltered, because this means no more than a replacement of some other solvent molecule in (a). Thus, clearly, the heat gained in the

---

[1] G. GEE and L. G. R. TRELOAR, *Trans. Faraday Soc.*, 38 (1942) 147.

[2] The values of $\triangle h_o$ were derived from the temperature coefficent of $\triangle g_o$ and their accuracy is not very high. The values obtained in later work are somewhat different from those quoted here, but they show a similar change of $\triangle h_o/\varphi_r^2$ with φ_r; see G. GEE and W. J. C. ORR, *Trans. Faraday Soc.*, 42 (1946) 507.

[3] For instance, A. E. VAN ARKEL, *Chem. Weekblad*, 31 (1934) 490; R. H. FOWLER and E. A. GUGGENHEIM, *Statistical Thermodynamics*, p. 357, Cambridge 1939.
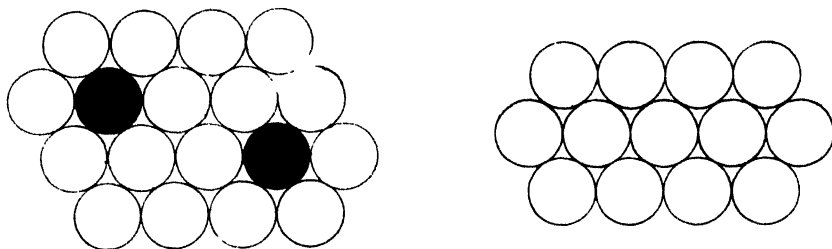
Fig. 8.   Transfer of solvent molecule from pure solvent (b) to solution (a).

process can differ from zero only in as far as configurations are concerned which
involve the interaction with at least two solute molecules. Since the number of such
interactions becomes *zero of a higher than the first order when the concentration becomes
zero*, the difference $\triangle h_o$ is *also small of a higher order than the concentration*. As shown
on p. 59, this accounts for the fact that RAOULT's and VAN 'T HOFF's law always
apply in the limit of zero concentration.

§ 6.   ENTROPY OF MIXTURES

a.   Entropy and probability

A given macroscopic state of a system may often be attained by means of a great
many different molecular arrangements. The "probability" W of a system with
given energy content is proportional to the number of distinct configurations which
are observant of the energy value concerned. It is shown in the statistical theory that
the entropy S of the system is related to the probability, W, by BOLTZMANN's equation

$$S = k \ln W \tag{38}$$

where k is BOLTZMANN's constant. This statistical interpretation of the entropy
plays an important part in modern science. The physical meaning of the second law
of thermodynamics may be compactly expressed by the following statement: con-
sidering a system at constant temperature, the equilibrium attained is the result
of a competition between the energy on the one hand, which tends to assume its
minimum value, and the entropy S on the other hand, which tends to reach its maximum
value. By way of illustration we may mention the equilibrium of a gas in a vessel
which is placed in a gravitational field. Here the lowest energy would correspond
to the situation where all gas molecules are lying at the bottom. The state of largest
entropy, however, is that where the gas molecules are distributed uniformly over
the entire vessel. The actual equilibrium represents a compromise between these two
conflicting tendencies. Another example is that of molecules carrying a permanent
dipole in a non-polar solvent. These molecules may show a pronounced tendency
to associate, thereby lowering their energy. The most probable state, however, is
that where all molecules are distributed and oriented at random. It would not be
difficult to add a great many further examples.

    In this connection a special case may be mentioned. In an ideal gas the energy
does not depend on the volume occupied. Consequently, the behaviour of the gas

at constant temperature is regulated by the entropy alone: the pressure p derives from the entropy ($p = T \, \partial S / \partial V$) and may be attributed entirely to the tendency of the molecules to acquire a uniform distribution in all space. Similarly, the osmotic pressure in an ideal solution (compare p. 133) may be said to result from the diffusion tendency of the molecules. In the chapter on random coils (p. 124) we shall meet with a still further example: the retractive force in stretched rubber. In all these cases the pressure (or the stress) is due to the fact that the system tends to attain the state of greatest probability.

b. Entropy of mixing in ideal solutions

b. 1. A central point in the discussion of solutions is the entropy gained at mixing. The nature of this entropy gain may be made clear as follows.

For definiteness let us call the component o the solvent, while the component 1 represents the solute. Let N_o molecules of the solvent be mixed with N_1 molecules of the solute. We assume that all molecules are of the same shape and size and that the number of nearest neighbours (coordination number) is equal for both kinds. Finally, to ensure that the simple interchange of any two molecules has no influence on the energy of the mixture, it is assumed that the interaction energies conform to the following equation:

$$a_{01} = \tfrac{1}{2} \, (a_{00} + a_{11})$$

It is at once obvious from equation (30) that this condition is sufficient, since in the present case the factor β in equation (31) becomes $\tfrac{1}{2} \, [a_{00} + a_{11} - 2a_{01}]$, and the heat of mixing is, therefore, zero.

Now let B be the number of possible arrangements in a liquid consisting of $N_0 + N_1$ *identical* molecules. The value of B is irrelevant to the present argument. If, instead of $N_0 + N_1$ identical molecules, we have N_0 molecules of component O and N_1 molecules of component 1, there will be $(N_0 + N_1) \, ! / N_0 \, ! \, N_1 \, !$ possible arrangements with the same energy content for each of the B arrangements mentioned. These are obtained by interchanging the $N_0 + N_1$ molecules among each other, taking into account that an interchange of *identical* molecules does not give rise to a new configuration. The total number of distinct configurations is, therefore

$$W = B \frac{(N_0 + N_1)!}{N_0 \, ! \, N_1 \, !} \tag{39}$$

Applying Boltzmann's equation (38) and using Stirling's approximation for $N!$ we obtain the entropy S of the mixture. We mention the final result only for the entropy of mixing:

$$S = - k \, [N_0 \ln(1 - x) + N_1 \ln x], \tag{40}$$

where x represents the mole fraction

$$x = \frac{N_1}{N_0 + N_1} \tag{41}$$

Differentiating with respect to N_o we obtain the *entropy of dilution* (per mole)

$$\triangle s_o = - R \ln (1 - x) \tag{42}$$

which is the result (24) used on p. 58 (R = gas constant).

It is important to realise that the logarithmic terms in equations (40) and (42) originate in the denominator in equation (39), i. e., *they are due to the fact that the interchange of identical molecules does not lead to a new configuration.* This logarithmic term in its turn gives rise to the well-known logarithmic term $RT \ln x$ in the molecular potential g_1 of the solute

$$g_1 = \left(\frac{\partial F}{\partial N_1}\right)_{V,T} = \left(\frac{\partial G}{\partial N_1}\right)_{p,T} = f_1 + RT \ln x \tag{43}$$

Similarly, the molecular potential of the solvent becomes

$$g_0 = f_0 + RT \ln (1 - x) \tag{44}$$

In these equations f_0 and f_1 are functions of pressure and temperature, but do not depend on x. Obviously f_0 is identical with the molecular potential g_0^0 of the pure solvent ($x = 0$). Solutions in which g_0 and g_1 are given by the simple expressions (43) and (44) are called ideal solutions. In these solutions the vapour pressure, osmotic pressure, etc. are determined solely by the entropy of mixing, i.e., by the tendency to attain the state of maximum probability.

b. 2. There actually do exist mixtures which show the properties of ideal solutions over a large range of concentrations [1]. The equations (43) and (44), however, owe their general importance to the fact that they govern the behaviour of all solutions provided these are sufficiently dilute. In fact, in non-ideal solutions we can no longer use the simple equation (39). Let us write

$$W = \frac{F(N_0, N_1)}{N_0! \, N_1!} \tag{45}$$

where F is some unknown function of the numbers N_0 and N_1. It is obvious that F will assume a *finite* value when N_1 approaches zero, since F will then simply approach the number of possible arrangements in the pure solvent. What is more, the manner in which this number is approached can also be indicated. For, if for a given configuration of the N_0 solvent molecules there are say B_1 possible arrangements for one single solute molecule (B_1 being a function of N_0), then, clearly, if N_1 is sufficiently small to neglect all interactions between solute molecules, the number of arrangements for N_1 solute molecules in this configuration is $B_1^{N_1}$ [2]. Thus it is obvious that not only F, but also $\ln F$ remains finite if N_1 approaches zero. The factor $N_1!$ in the denominator of (45) must be maintained under all circumstances to express that an interchange of identical molecules gives no new configuration. From $N_1!$ we derive the logarithmic term $RT \ln x$ in the equation for g_1. Instead of (43) the expression for g_1 will now contain additional terms which depend on x, but *if x is sufficiently small the logarithmic term will always predominate over all other terms.* This means that the solution shows ideal properties when it is sufficiently dilute. The argument given here would not apply if it were possible that the numerator F of W contained a factor which tended to zero when N_1 approached zero. We know, however, that F and $\ln F$ remain finite. The pregnant fact that g_1 becomes logarithmically infinite in very dilute solutions has a simple physical reason: the work done by a mole of

[1] J. H. HILDEBRAND, *Trans. Faraday Soc.*, 33 (1937) 144; *J. phys. Chem.*, 43 (1939) 109, 297.
[2] A. J. STAVERMAN, *Rec. trav. chim.*, 60 (1941) 76.

the solute when being distributed over an infinite amount of solvent is infinite, just as is the work done in expanding a mole of a gas to an infinite volume. From the relation (17) on p. 57 we may derive that the molar free energy g_0 of the solvent, when expanded in powers of x, contains the term $- RTx$, which accounts for RAOULT's law and VAN 'T HOFF's law in dilute solutions[1]. The result obtained is quite general and applies to molecules of any size and shape, including macromolecules dissolved in a micromolecular solvent. The only difference is, that extremely low concentrations must be attained to verify the laws of ideal solutions (RAOULT's law, VAN 'T HOFF's law). This has sometimes led to the conclusion that the properties of macromolecules in solution are not in conformity with these laws[2].

Since the question is one of fundamental character, the following illustration may be added, which at the same time may serve as an introduction to the theory developed on p. 72.

It is obvious that the number of arrangements in a solution of macromolecules will be much larger than that of simple spheres, especially if the chains may adopt a large number of different shapes. This, however, plays no part in the logarithmic term $RT \ln x$, as is perhaps best shown by the following simple example. Let the solvent molecules be spheres, denoted by the index o, while the solute molecules are dumb-bells consisting of two spheres 1 and 1* which are of the same size as the solvent molecules[3]. For definiteness let the coordination number be 12. For the first solute molecule the sphere 1 may be placed on $N_0 + 2N_1$ lattice points. This can be done in 12 ways, because the sphere 1* may occupy any of the twelve neighbouring sites. The second solute molecule can then be placed in $12(N_0 + 2N_1 - 2)$ ways, provided the solution

Fig. 9.   Dumb-bells dispersed among spheres.

is sufficiently dilute to ignore the number of arrangements in which the first molecule interferes with the position of the second. In this way we find a total number of

$$12^{N_1} \, 2^{N_1} \frac{(\tfrac{1}{2}N_0 + N_1)!}{(\tfrac{1}{2}N_0)!}$$

possible arrangements for the solute molecules. Finally, the N_0 solvent molecules can be placed on the N_0 remaining lattice sites in $N_0!$ ways. To find the total number of *distinct* configurations we must divide by $N_0! N_1!$. Thus

$$W = 12^{N_1} \, 2^{N_1} \frac{(\tfrac{1}{2}N_0 + N_1)!}{(\tfrac{1}{2}N_0)! \, (N_1)!}$$

This leads to the following simple result for the molar free energy of the dumb-bells:

$$g_1 = RT \left[\ln \frac{N_1}{N_0 + 2N_1} - \ln 12 \right] = RT \left[\ln \frac{x}{1 + x} - \ln 12 \right]$$

Here again, in the limit $x \to 0$ the terms $RT \ln x$ predominates. *The increase in the number of possible arrangements does evidently not affect this term, but only introduces the term $- RT \ln 12$ in the expression for g_1.* It is clear that a similar argument applies to longchain molecules consisting of a large number of spheres. Deviations will only occur at higher concentrations where it is no longer permissible to neglect the interference of a solute molecule with the possible positions of other solute molecules (compare p. 73).

---

[1] J. J. HERMANS, *Rec. trav. chim.*, 60 (1941) 370.
[2] Critical discussion in E. HÜCKEL, *Z. Elektr.*, 42 (1936) 753.
[3] R. H. FOWLER and G. S. RUSHBROOKE, *Trans. Faraday Soc.*, 33 (1937) 1272.

b. 3. The result obtained regarding the ideal character of all solutions in the limit of small concentrations may be expressed by saying that the free energy of dilution is of the form: $\triangle g_0 = -RTx +$ terms of a higher order than the first in x. If then we apply the equation

$$\triangle h_0 = \frac{\partial \triangle g_0 / T}{\partial 1/T}$$

for the heat of dilution $\triangle h_0$ we find immediately that this heat of dilution becomes of a higher order than the first when x approaches zero. The significance of this result was discussed on p. 59, and a molecular interpretation was given on p. 65.

c. Equilibrium between polymer and ideal solution

The discussion of real (i.e., non-ideal) polymer solutions will be deferred till section 6 d. We shall first consider the "solubility" of macromolecules from the point of view of ideal solutions. This aspect has played an important part in the physical chemistry of high polymers [1], especially in the theory of fractional precipitation.

At one time it was believed that the entropy of polymer solutions followed approximately the ideal laws [2], and the non ideal character was attributed to energy effects. Consider now the distribution of macromolecules over two liquid phases, for instance two non-miscible liquids. Designating one of the phases by a stroke, and introducing the subscript r for the polymer molecules, the equilibrium requires that

$$h_r - TS_r = h'_r - Ts'_r$$

If the entropy follows the ideal law, we may write

$$s_r = \sigma_r - R \ln x; \quad s'_r = \sigma'_r - R \ln x' \tag{46}$$

where σ_r and σ'_r are independent of the concentrations. This gives

$$\frac{x}{x'} = e^{\frac{\sigma_r - \sigma'_r}{R}} \; e^{-\frac{h_r - h'_r}{RT}} \tag{47}$$

If the ideal law (46) is meant to apply to the entire concentration range, the entropy constants σ_r and σ'_r are equal, because in that case they simply represent the molar entropy of the pure polymer (x resp. x' equal to 1). Writing $E = h_r - h'_r$, the equation (47) then assumes the familiar form

$$\frac{x}{x'} = e^{-\frac{E}{RT}} \tag{48}$$

With linear long-chain molecules, whether straight, or loosely coiled, the energy difference E will be approximately proportional to the degree of polymerisation P,

[1] J. N. BRØ NSTED, *Z. physik Chem* , A. Bodenstein Band, (1931) 257. J. N. BRØNSTED and E ARNING, *Z physik. Chem.*, A 155 (1931) 343; G. V. SCHULZ, *Z. physik. Chemie*, A 179 (1937) 321; W4 (1939) 1; B 46 (1940) 105, 137; *J. prakt. Chemie*, 155 (1940) 115.

18   [2] G. V. SCHULZ, *Z. physik. Chem.*, B 52 (1942) 253.

or, to use the variable introduced in the present chapter, the number of segments n. In which case

$$\frac{x}{x'} = (e^{-\frac{\epsilon}{RT}})^n = (\text{const})^n \tag{49}$$

This shows at once that if $\epsilon > 0$ (const. < 1), the ratio x/x' will be very small if n is large. One of the phases will then always be very poor in polymer substance. In particular, when we consider the equilibrium between the pure polymer ($x' = 1$) and the solution, we will find that x is extremely small: the polymer is *practically insoluble*. This explains why it was believed that macromolecules could be dissolved only when heat was evolved on mixing (compare p. 61). As soon as ϵ is negative (const. > 1), the polymer shows *complete miscibility*. Intermediate cases of limited solubility do not exist[1], unless ϵ happens to be very small (critical temperature or critical composition of the solvent). With compact molecules, e.g., globular proteins, E will be proportional to the surface of the particle rather than to its mass, but qualitatively the result will be exactly the same.

Although this theory accounts for a very important phenomenon in the physical chemistry of macromolecules, it fails utterly in those cases where dissolution occurs in spite of a positive E-value. Such cases are reported in the physical chemistry of rubbers[2] and can only be explained by the non-ideal entropy of mixing (see section 6. d). Moreover, the experiments show that the polymer usually is not really insoluble in those cases where there does not exist complete miscibility: the equilibrium attained is not an equilibrium between polymer and extremely dilute solution, but between a concentrated and a dilute phase. The composition of the concentrated phase is independent of the degree of polymerization, while that of the dilute phase is the smaller the larger the molecular weight. This will be explained on p. 78.

Yet we should not reject the theory entirely on these grounds. When the heat of mixing (either positive or negative) is very large, it becomes the determining factor, and in that case the formula (48) will at least qualitatively lead to the correct result. It was pointed out by FRUMKIN[3], that this formula explains the distribution of the lower homologous alcohols (up to 5 C-atoms) in the system benzene-water or water-cotton seed oil. Here the deviations from the ideal entropic behaviour are less pronounced, and the energy E in formula (48) is increased by a constant amount at the addition of each new C-atom in the chain. A similar result was obtained by LANGMUIR[4] as early as 1917 in a discussion of surface adsorption. Here the application of the formula (48) to the equilibrium between the surface layer and the bulk of a liquid leads to TRAUBE's rule on the change of surface activity in a homologous series. Similarly, the vapour pressure of normal paraffins decreases exponentially with increasing chain-length[5], and so on. With increasing molecular weight of the solute, however, the deviations from ideal entropic behaviour become progressively more pronounced. These will be the subject of the ensuing section.

[1] BRØNSTED, *loc. cit.*
[2] G. GEE, *Trans. Faraday Soc.*, 38 (1942) 276, 418; 40 (1944) 463; 41 (1945) 340; *Trans. I. R. I.*, 18 (1943) 266.
[3] A. FRUMKIN, *Z. phys. Chem.*, 116 (1925) 501.
[4] I. LANGMUIR, *J. Am. Chem. Soc.*, 39 (1917) 1894.
[5] BRØNSTED, *loc. cit.*

d. Entropy in non-ideal polymer solutions

d. 1. Introductory remarks

It was pointed out in section 6 *b*. 2 (small print, p. 69) that the number of configurations in polymer solutions will be far greater than that in solutions of simple molecules. Qualitatively this question has been considered already by MEYER [1]. Attempts to obtain a quantitative solution of the problem were made by STAVERMAN [2] and by MÜNSTER [3]. The crucial point in these calculations is, of course, to account for those molecular arrangements in which a solute molecule interferes with the position of other solute molecules. A satisfactory solution of this difficulty was given by HUGGINS [4] and by FLORY [5]. MILLER [6] applied an inductive method: he calculated the entropy of a mixture containing spheres (solvent molecules) and three-atomic solute molecules, each consisting of three spheres. Comparing his result with the one obtained by FOWLER and RUSHBROOKE [7] for dumb-bells immersed in spheres, he was able to generalise the formulae for arbitrary chain-lengths. MILLER's result is essentially the same as that found in HUGGINS' or FLORY's theory, of which we shall now give a very short account, restricting ourselves to the principles of the method.

d. 2. Calculation of entropy

We assume the model introduced on p. 60. Each polymer molecule may consist of n segments; a segment replaces a solvent molecule on the quasi lattice of the liquid. If there are N_o solvent molecules and N_r polymer molecules, the volume fractions are (equations 29)

$$\varphi_o = \frac{N_o}{N_o + nN_r} \; ; \quad \varphi_r = \frac{nN_r}{N_o + nN_r} \tag{50}$$

For the present we assume zero heat of mixing; the influence of energy effects will be discussed on p. 76. We shall estimate the number of possible arrangements by first placing all polymer segments on lattice sites and then filling up the remaining lattice sites with solvent molecules. As there are $N_o + nN_r$ sites, one end of the first chain molecule can be placed in $N_o + nN_r$ ways. If the coordination number of the lattice is z, the second segment of this chain can be placed in z ways. For the third segment the number of possible positions is less than z, say y, since one of the z lattice points surrounding the second segment is already occupied by the first segment, and, moreover, the chains are not completely flexible; with rigid rods we would have $y = 1$. Now, the 4th, 5th, etc. segment can also be placed in y ways, provided we may neglect the small fraction of positions which have already been occupied by segments of the same chain. The total number of possible arrangements for the first chain-molecule is therefore $(N_o + nN_r)zy^{n-2}$.

---

[1] K. H. MEYER, *Hochpolymere Chemie*, II, p. 549, Leipzig 1940.

[2] A. J. STAVERMAN and J. H. VAN SANTEN, *Rec. trav. chim.*, 60 (1941) 76; A. J. STAVERMAN, *Rec. trav. chim.* 60 (1941) 640.

[3] A. MÜNSTER, *Kolloid-Z.*, 105 (1943) 1.

[4] M. L. HUGGINS, *J. phys. chem.*, 46 (1942) 151; *J. Am. Chem. Soc.*, 64 (1942) 1712; *Ind. Eng. Chem.*, 35 (1943) 216.

[5] P. J. FLORY, *J. chem. Physics*, 9 (1941) 660; 10 (1942) 51; 12 (1944) 425.

[6] A. R. MILLER, *Proc. Cambr. Phil. Soc.*, 38 (1941) 109; 39 (1942) 53.

[7] R. H. FOWLER and G. S. RUSHBROOKE, *Trans. Faraday Soc.*, 33 (1937) 1272.

The number of lattice points available to the first segment of the second polymer molecule is $N_o + nN_r - n$. The number of possible positions for the second one is *less than* z, because some sites are already occupied by the first polymer molecule. We replace z by $z(1 - f_2)$ and assume that f_2 is approximately equal to the fraction $n(N_o + nN_r)^{-1}$ of occupied sites. For the same reason the number of possible positions of the 3 d, 4 th, etc. segment will no longer be y but $y(1 - f_2)$. Consequently, this second polymer molecule can be placed in $(N_o + nN_r - n)zy^{n-2}(1 - f_2)^{n-1}$ ways.

For the third polymer molecule we will find $(N_o + nN_r - 2n)zy^{n-2}(1 - f_3)^{n-1}$, where f_3 is approximately equal to the fraction $2n(N_o + nN_r)^{-1}$ of sites which are occupied by the first and second polymer molecule. And so on. Finally, the N_o solvent molecules can be placed on the N_o remaining lattice sites in N_o! ways. To find the number of *distinct* configurations we will have to divide by $N_o! N_r!$ and so finally obtain [1]

$$W = \frac{n^{N_r}}{N_r!} \frac{(N_o/n + N_r)!}{(N_o/n)!} z^{N_r} y^{(n-2)N_r} (\text{-})^{(n-1)} \tag{51}$$

$$(\text{-}) = (1 - f_2)(1 - f_3) \ldots (1 - f_{N_r}), \text{ and } f_i = \frac{(i-1)n}{N_o + nN_r} \tag{52}$$

The factor (-) is essential for the influence of the solute molecules on each other's possible positions. The value of (-), or rather of $ln(\text{-})$ may be approximated by replacing $\Sigma_i \ln (1 - f_i)$ by an integral:

$$\ln (\text{-}) \simeq \int_0^{N_r} ds \ln(1 - \frac{sn}{N_o + nN_r}) = -N_r - \frac{N_o}{n} \ln \frac{N_o}{N_o + nN_r}$$

Using STERLING's formula for $N!$ we finally obtain

$$\ln W = N_r [\ln(nzy^{n-2}) - (n-1)] + N_o \ln \frac{N_o + nN_r}{N_o} + N_r \ln \frac{N_o + nN_r}{nN_r} \tag{53}$$

In the pure (amorphous) polymer we have $N_o = 0$, and the entropy is

$$S = k \ln W = N_r k[\ln(nzy^{n-2}) - (n-1)] \tag{54}$$

This shows that the entropy of an amorphous macromolecular substance depends on the number n, i.e., on the chain-length. This result was used by FRITH and TUCKETT [2] to explain the comparatively long melting interval (compare p. 55). Returning to formula (53), we observe that this expression becomes zero in the limit $N_o = 0$. *The entropy of mixing* $\triangle S$ is the difference between the entropy of the mixture and that of the pure components. Using the volume fractions we find therefore

$$\triangle S = - kN_o \ln \varphi_o - k N_r \ln \varphi_r \tag{55}$$

---

[1] Symmetry factors are omitted here, because they are irrelevant to the present argument. They are equally effective in the mixture and in the pure components, and therefore do not occur in the final result (55) for the entropy of mixing, although they should be included in the formula (54) for the entropy of the pure polymer. With linear long-chain molecules of which the two ends are equal, the symmetry number is 2, which means that a term $- N_r \ln 2$ should be added to (54).

[2] E. M. FRITH and R. F. TUCKETT, *Trans. Faraday Soc.*, 40 (1944) 251.

Comparing with the ideal entropy of mixing (equation 40 on p. 67), i.e.

$$S = - k \, N_o \ln (1 - x) - k \, N_r \ln x,$$

we observe that the mole fractions are replaced by the volume fractions.

d. 3. Dilute solutions

Applications of the result obtained will be discussed at length in the ensuing sections. At present we shall go no further than to indicate its importance to the physics of *dilute* polymer solutions. Differentiating the formula (55) with respect to N_o we obtain the entropy of dilution

$$\triangle s_o = - R \ln q_o - R(1 - \frac{1}{n})q_r \tag{56}$$

In the limit of infinite dilution this conforms to the ideal law. In fact, when expanding in powers of volume fraction q_r, and mole fraction x, respectively,

$$\frac{\triangle s_o}{R} = \frac{1}{n} q_r + \tfrac{1}{2} q_r{}^2 + \tfrac{1}{3} q_r{}^3 \ldots = x + \tfrac{1}{2}(n^2 - 2n + 2)x^2 \tag{57}$$

The first power of x in this series is also obtained in the expansion of the ideal expression $\triangle s_o = - R \ln (1 - x)$, showing that in the limit of infinite dilution RAOULT's and VAN 'T HOFF's law will apply. It is obvious from equation (57), however, that deviations from these laws will become apparent at very small concentrations because n is a large number. In Fig. 10 the vapour pressure calculated from equation (23), p. 58, ($RT \ln p/p_o = - T \triangle s_o$, since $\triangle h_o$ is here assumed to be zero) is plotted against mole fraction for various values of n. In the limit $x = 0$ all curves are tangential to RAOULT's straight line for $n = 1$, but obviously with high values of n this can only be verified at extreme dilutions. Further it is obvious from equation (57) that it is permissible to expand no further then to the second power of q_r if $q_r \ll 1$. This second power cannot be neglected compared with the first unless $nq_r \ll 1$. This result will be applied in section 8 a, p. 85,
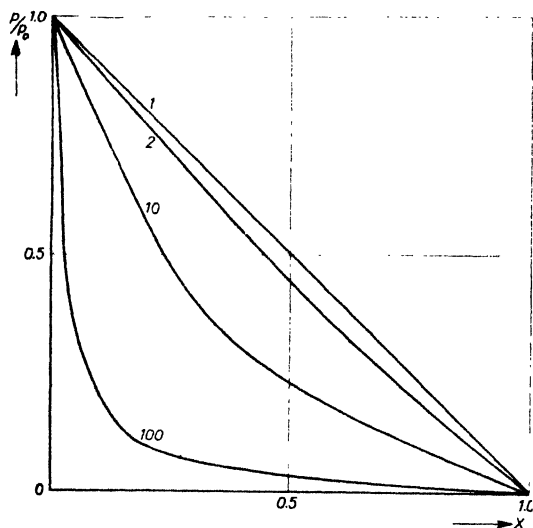
Fig. 10. Vapour pressure of solvent for various values of n (zero heat of mixing).

where the osmotic pressure is developed in a power series, neglecting terms of the order $q_r{}^3$.

d. 4. Extremely dilute solutions

To conclude, it should be pointed out, that the theory sketched here tacitly assumes that the concentration of the polymer component is not too small. In extreme dilution each polymer molecule forms practically an isolated system, and the probability for a given lattice site to be occupied by a polymer segment becomes strongly dependent on whether or not adjacent lattice sites are occupied by a polymer segment. The introduction of the overall factors $1 - f_2$, $1 - f_3$ etc. to account for the influence of a solute molecule on the possible positions of other solute molecules must then be replaced by a more refined treatment. FLORY[1] has considered this effect in some detail; the correction involved is in conformity with experimental results at great dilution.

e. Thermodynamic behaviour of a polymer homologous mixture

We are now in a position to consider the melting process in a polymer homologous mixture. It was pointed out on p. 53 that the melting point of such a mixture need not necessarily be the same as that of the components, even though the pure components all have the same melting point. We shall show now that it is only under special circumstances that the melting point of a polymer is not affected by the addition of a homologue.

Consider the pure component 1 at its melting point. Denoting the solid phase by a stroke, we have $f_1 = f_1'$ ($f =$ molar free energy divided by RT). Since the second component has the same melting point we have at this temperature also: $f_2 = f_2'$. Let n_1 and n_2 respectively be the numbers of segments in the molecules. At the addition of a small amount of the second component, containing n_2 segments per molecule, we obtain a mixture whose entropy will be, according to result (55), $-kN_1 \ln \varphi_1 - kN_2 \ln \varphi_2$, where the volume fractions φ_1 and φ_2 are determined by

$$\varphi_1 = \frac{n_1 N_1}{n_1 N_1 + n_2 N_2} \qquad \varphi_2 = \frac{n_2 N_2}{n_1 N_1 + n_2 N_2}$$

The molar free energy (divided by RT) of component 1 in the liquid phase is thereby changed from f_1 tot $f_1 + \ln \varphi_1 + (1 - n_1/n_2)\varphi_2$, and that of component 2 to $f_2 + \ln \varphi_2 + (1 - n_2/n_1)\varphi_1$. It depends on the solid phase whether this change in the liquid state will leave the equilibrium unaffected. If the solid is an amorphous mixture, whose thermodynamic properties depend on the composition in exactly the same way as those of the liquid phase, the equilibrium requires that

$$f_1 + \ln \varphi_1 + (1 - \frac{n_1}{n_2})\varphi_2 = f_1' + \ln \varphi_1' + (1 - \frac{n_1}{n_2})\varphi_2'$$

$$f_2 + \ln \varphi_2 + (1 - \frac{n_2}{n_1})\varphi_1 = f_2' + \ln \varphi_2' + (1 - \frac{n_2}{n_1})\varphi_1'$$

These equations have the solution $\varphi_1 = \varphi_1'$ (i.e., $\varphi_2 = \varphi_2'$) with $f_1 = f_1'$ and $f_2 = f_2'$. This means that the equilibrium is maintained at the same temperature; the composition of the solid phase is the same as that of the liquid phase. In other words, the mixture behaves as a one-component system, which melts without change in composition and whose melting point is independent of the composition. In general, however, the free energy of the solid will not be given by exactly the same expression as that of the liquid. For example, if the entropy in the solid were approximately ideal, we should have

$$f_1 + \ln \varphi_1 + (1 - \frac{n_1}{n_2})\varphi_2 = f_1' + \ln(1 - x')$$

$$f_2 + \ln \varphi_2 + (1 - \frac{n_2}{n_1})\varphi_1 = f_2' + \ln x'$$

[1] P. J. FLORY, J. chem. Phys., 13 (1945) 453.

where x' is the mole fraction $N'_2/(N'_1 + N'_2)$ in the solid. For these two equations to apply simult-aneously, the equalities $f_1 - f'_1$ and $f_2 = f'_2$ can no longer be maintained, which means that the temperature will differ from the melting points of the pure components.

It is further obvious from these considerations that the equilibrium between a polymer homo-logous mixture and a solution of this mixture in a solvent of low molecular weight is a very intricate affair. The composition of the polymer homologous mixture in the solid, (which in most cases is a gel containing a certain amount of the solvent), is different from that in the solution. At the addition of further amounts of polymer substance, the equilibrium is disturbed, in contrast to the solubility equilibrium in micromolecular substances. This is the so-called ,,BODENKÖRPER'' rule, according to which the solubility of a polymer depends on the total amount of polymer present. This rule is only apparent, however, since no such effect exists in the solubility of a well-fractionated sample.

Further details concerning the thermodynamics of a polymer homologous mixture can be found in papers by FLORY[1] and SCOTT[2].

§ 7. FREE ENERGY OF POLYMER SOLUTIONS

The theory of the preceding section assumes a completely random distribution of polymer segments. Strictly speaking, this is not permissible when forces act. If heat is absorbed on mixing, polymer-polymer contacts and solvent-solvent contacts will be relatively more frequent than polymer-solvent contacts. Instead of calculating the entropy and the heat content on the assumption of random mixing, one should start from the so-called phase integral (partition function)[3]. However, when the heat of mixing is small, the result obtained is the same as if mixing were random, as would be expected. Simple results in explicit form have not yet been obtained for more general cases, and for the present we prefer the simplified treatment to the more general one, the more so as we are only interested in the general character of the results obtained. Considerations on the basis of the phase-integral may be found in the work of GUGGENHEIM[4], ORR[5], and MILLER[6].

Since our model assumes that there is no appreciable volume change on mixing, we need not distinguish between the free energy and the GIBBS free energy of the liquid mixture. The GIBBS free energy of mixing $\triangle G = \triangle H - T \triangle S$ becomes, according to equations (31), p. 63, and (55), p. 73.

$$\triangle G = kT(N_o + nN_r) \ (q_o \ln q_o + \frac{1}{n} q_r \ln q_r + \gamma q_o q_r) \tag{58}$$

where

$$\beta = kT\gamma \tag{59}$$

The constant γ is, therefore, determined by the interaction energies (see p. 62). Negative values of γ indicate attraction between polymer and solvent, and vice versa. In the following it is to be borne in mind that the constant γ depends on the tem-perature, whereas the accompanying logarithmic terms in (58) are independent of T (see further p. 82).

[1] P. J. FLORY, *J. Chem. Phys.*, 12 (1944) 425.

[2] R. L. SCOTT, *J. Chem. Phys.*, 13 (1945) 178.

[3] R. H. FOWLER, *Statistical Mechanics*, R. H. FOWLER and E. A. GUGGENHEIM, *Statistical Thermodynamics*, Cambridge 1939.

[4] E. A. GUGGENHEIM, *Proc. Roy. Soc.*, A 183 (1944) 213; *Trans. Faraday Soc.*, 41 (1945) 107.

[5] W. J. C. ORR, *Trans. Faraday Soc.*, 40 (1944) 306.

[6] A. R. MILLER, *Proc. Cambr. Phil. Soc.*, 38 (1941) 109; 39 (1942) 53.

a. Separation into two layers; solubility of high polymers

a. 1. Equilibrium between dilute and concentrated phase

In Fig. 11 $\triangle G$ is plotted against volume fraction φ_r for various values of γ and n. The values of $\triangle G$ in this figure refer to a constant volume of the mixture, i.e., $N_o + nN_r = constant$.
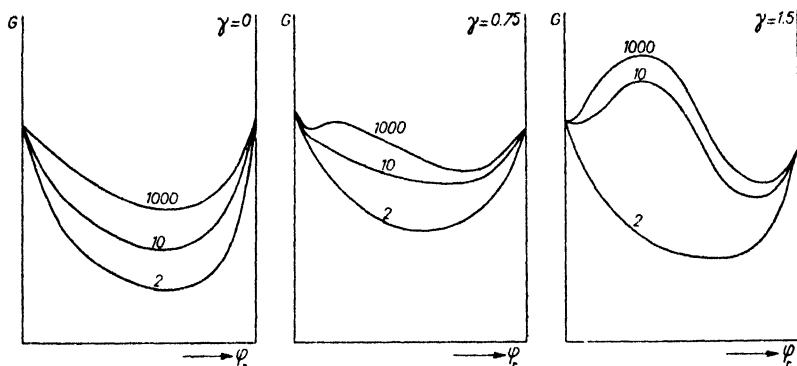
Fig. 11. Gibbs free energy of mixing against volume fraction of polymer
for various values of γ and n.

From a critical value of γ onwards the curves for $\triangle G$ show two minimum values. Any solution whose composition is represented by a point between the two tangent points t end t' (miscibility gap) will separate spontaneously into two layers, the one rich and the other poor in polymer substance. There is nothing remarkable in this result when γ has a large positive value. In polymer solutions, however, these minimum values occur at γ-values which are considerably smaller than in solutions of micromolecules.

Further, as shown in the figures, the amount of polymer substance in the dilute phase will be the smaller the larger is n (i.e., the larger the degree of polymerisation). For the sake of clarity, we even had to exaggerate the Figure 11, since in reality the minimum at the left occurs at concentrations φ_r which are smaller than those shown in the figure. If the concentrated phase is distinguished by strokes, the equilibrium conditions are

$$\triangle g_o = \triangle g'_o \; ; \; \triangle g_r = \triangle g'_r \tag{60}$$

Now, from equation (58) we derive

$$\frac{\triangle g_o}{RT} = \ln \varphi_o + (1 - \frac{1}{n})\varphi_r + \gamma \varphi^2_r \tag{61}$$

$$\frac{\triangle g_r}{RT} = \ln \varphi_r + (1 - n) \varphi_o + \gamma n \varphi^2_o \tag{62}$$

It is interesting to observe that these equations are completely *symmetrical*. To that end we remember that n is equal to the ratio v_r/v_o of the molar volumes of polymer and solvent, and further introduce the constant $\varkappa = \gamma/v_o$. This gives

$$\frac{\triangle g_o}{RT} = \ln\varphi_o + (1 - \frac{v_o}{v_r})\varphi_r + v_o\,\varkappa\varphi^2{}_r$$

$$\frac{\triangle g_r}{RT} = \ln\varphi_r + (1 - \frac{v_r}{v_o})\varphi_o + v_r\,\varkappa\varphi^2{}_o$$

Using the equations (61) and (62) in the equilibrium conditions (60) we find

$$\ln\varphi'_o + (1 - \frac{1}{n})\varphi_r + \gamma\varphi^2{}_r = \ln\varphi'_o + (1 - \frac{1}{n})\varphi'_r + \gamma\varphi'^2{}_r \qquad (63)$$

$$\ln\varphi_r + (1 - n)\varphi_o + \gamma n\varphi^2{}_o = \ln\varphi'_r + (1 - n)\varphi'_o + \gamma n\varphi'^2{}_o \qquad (64)$$

These equations determine the volume concentrations φ_r and φ'_r of the polymer in the two phases. Now, n is a large number, and we know from Fig. 11 that φ_r is very small ($\varphi_o \cong 1$). The equilibrium condition (63) therefore becomes, practically, $\ln(1 - \varphi'_r) + \varphi'_r + \gamma\varphi'^2{}_r = 0$. Its solution is evidently independent of n.

This explains why the volume concentration in the concentrated phase (for a given polymer substance, in a given solvent, at a given temperature) *is practically independent of the molecular weight of the polymer.* Further, since φ_r is small and n large, equation (64) may be approximated to

$$\ln\varphi_r - (1 - \gamma)n = - n\varphi'_o + n\gamma\varphi'_o{}^2, \quad \text{or}$$

$$\ln\varphi_r = n\varphi'_r\,[1 - \gamma(2 - \varphi'_r)] \qquad (65)$$

We know already that φ'_r is independent of n and thus (65) has exactly the same form as (49):

$$\varphi_r = (\text{const})^n \qquad (66)$$

The constant in this equation is > 1 when $1 - \gamma(2 - \varphi'_r) > 0$ (in which case there is complete miscibility) and < 1 when $1 - \gamma(2 - \varphi'_r) < 0$. Obviously the smallest value of γ, where a separation into two layers is at all possible is: $\gamma = \frac{1}{2}$. As pointed out already on p. 71 the result (66) indicates that the polymer is either completely miscible with the solvent or practically insoluble. This term "insoluble", as we know now, refers to the dilute phase; it does by no means imply that the polymer is not dissolved; on the contrary, the "concentrated" phase itself may be rather poor in polymer. It is obvious that these results are of paramount importance to the theory of *coacervates*, discussed in Chapter X (see also p. 90 in this chapter).

a. 2. Solubility

The equation (62) may further be applied to the equilibrium between pure polymer and solution [1]. The GIBBS potential g'_r of the pure polymer is approximately proportional to the chain length (see p. 52). The equilibrium condition therefore assumes the form

[1] M. L. HUGGINS, *J. Am. Chem. Soc.*, 64 (1942) 1717; *J. Phys. Chem.*, 46 (1942) 156.

$$\ln\varphi_r + (1 - n)\varphi_o + \gamma n\varphi^2_o = nK,$$

where K depends on the temperature but is independent of n. If n is large and the polymer not very soluble, this gives

$$\ln\varphi_r = -n\,(\gamma - 1 - K) \qquad (67)$$

which is again of the form (66). For shorter chain-lengths an equation may be derived involving the heat of fusion, which may be derived from experiment. On this basis, HUGGINS [1] calculated the solubility of normal paraffins in decalin as a function of temperature and found complete agreement with experiment (Fig. 12).

As pointed out on p. 75, complications arise when the polymer is not fractionated. According to RICHARDS [2] further

Fig. 12. Solubility (volume fraction) of n $C_{34}H_{70}$ and n $C_{60}H_{122}$ in decalin, as a function of temperature. Curves theoretical; points experimental.

complications may be due to the composite nature of the solid phase, because in most cases this solid phase is a mixture of crystalline and amorphous material.

b. Fractional precipitation

b. 1. *Addition of non-solvent*

When changing the temperature or the composition of the medium, the value of VAN LAAR's constant is changed. For example, at the addition of a non-solvent to the solution of a polymer substance, γ is increased and, at a certain critical value of γ, separation occurs. Under these circumstances it is no longer valid, that the polymer is "either insoluble or completely miscible"; there is a finite solubility which depends on the degree of polymerization. The larger the molecular weight of the polymer, the sooner its solubility limit is reached. In this way it is possible by progressive addition of non-solvent to precipitate successive fractions of the macro-molecular substance.

The effect can be accounted for approximately on the basis of SCHULZ's theory [3], which was developed for systems obeying the relation (48) p. 70, but which can also be made to apply to the present theory. To that end we assume that the value of γ which determines the interaction of polymer and solvent is a linear function of the composition of the solvent (SCHULZ made the practically equivalent assumption that E in equation 48 is a linear function of this composition). If the changes in

[1] M. L. HUGGINS, *J. Am. Chem. Soc.*, 64 (1942) 1717; *J. Phys. Chem.*, 46, (1942) 156.

[2] R. B. RICHARDS, *Trans. Faraday Soc*, 42 (1946) 10, 20.

[3] G. V. SCHULZ, *Z. physik. Chem.*, A 179 (1937) 321; 184 (1939) 1; *J. prakt. Chem.* 155 (1940) 115; *Z. physik. Chem.*, B 46 (1940) 105, 137.

composition to be considered are not too large, this approximation is reasonable.
If C is the concentration of the non-solvent, we may, therefore, write

$$\gamma = a + bC \tag{68}$$

As a measure for the "precipitability" of a polymer we consider the critical concentration C^* of the non-solvent which is needed to reduce the solubility (expressed in volume fraction) to a certain critical (low) value φ^*_r. . If the polymer is precipitated in a pure undiluted state, it is at once obvious from equation (67) that the critical concentration will be reached the sooner the larger is n, i.e., the larger the degree of polymerisation P. We derive from this equation that

$$C^* = AP^{-1} + B \tag{69}$$

where A and B are constants. If the polymer molecules are compact structures (globular proteins or strongly branched chain molecules) the precipitation is caused by a surface action which is proportional to $P^{2/3}$ rather than P. In these cases we should expect a relation of the type

$$C^* = AP^{-2/3} + B \tag{70}$$

Most experiments are in conformity with an equation of the type $C^* = AP^{-s} + B$, but the range of P-values which is accessible to accurate experiment is usually insufficient to ascertain the value of s in a decisive manner (compare Fig. 5 on p. 144 in the chapter on molecular weight determinations).

In most cases the polymer is precipitated in the gel state; i. e., we are dealing with a separation into two mixtures to which the theory of section 7. a. 1 should be applied. The original solution is supposed to have a γ-value below $\frac{1}{2}$, which implies complete miscibility. With increasing concentration C of non-solvent, γ is increased until at a critical concentration C^* a turbidity is observed, indicative of separation. This occurs the sooner, the higher the molecular weight of the polymer. It can be shown from the equations (61) (62) and (68) that here again the concentration C^* conforms approximately to an equation of type (69). In the case of a polymer homologous mixture, the fractions with the highest molecular weight are precipitated first; i. e., the concentrated phase is particularly rich in the high fractions. Some experimental aspects of the method are discussed on p. 144 (chapter on molecular weight determinations).

b. 2. Change of temperature

In order to achieve a change in γ one may also change the temperature of the mixture. According to HUGGINS, γ is approximately proportional to the reciprocal absolute temperature (see p. 82).

Some aspects of this problem were discussed bij BRÖNSTED and WARNING [1] as early as 1931. They considered a solvent mixture at a critical temperature T_k (Fig. 13). Let the composition be determined by the mole fraction x. When the temperature T drops slightly below T_k, the solvent separates into two layers of compositions x_1 and x_2. For small values of $\varDelta T = T_k - T$, any curve (Fig. 13) can be approximated

[1] J. N. BRÖNSTED and E. WARNING, Z. Physik. Chem., A 155 (1931) 343. See further the interesting work of J. N. BRÖNSTED and K. VOLQUARTZ, Trans. Faraday Soc., 36 (1940) 619.

by a parabola $(x - x_k)^2 = a^2(T_k - T)$, where a is a constant. The two roots x_1 and x_2 of this equation are at equal distances from x_k :

$$\Delta x = x_k - x_1 = x_2 - x_k = a\sqrt{\Delta T}$$

Suppose now, that a third substance is dissolved in this mixture, and let Ψ be the distribution ratio of this substance over the two phases. The value of Ψ depends on the composition x and on the temperature T. Since $\Delta T/\Delta x = 0$ at the critical temperature, we may write $\Delta \Psi = (\partial \Psi/\partial x)_{xk} \Delta x$ and therefore $\Delta \Psi = b \sqrt{\Delta T}$, where b is a constant. In particular, when the dissolved substance is a colloid, the value of b is very large and the distribution ratio becomes strongly different from unity even at very small values of ΔT: one of the phases becomes very rich, the other very poor in colloid. Colorimetric experiments with As_2S_3 and Cr_2O_3 sols in mixtures of butyl alcohol and ethyl alcohol $(T_k = 20°)$ have confirmed the proportionality between $\Delta \Psi$ and $\sqrt{\Delta T}$.

In these considerations the solute practically plays no part in the thermodynamic properties of the system. When separation occurs in polymer

Fig. 13. Two component system at critical temperature.

solutions, however, the high polymer substance itself is determinant for the thermodynamic behaviour and we must apply the theory of section 7. a. 1, taking into account that the interaction constant γ changes with temperature (see below). Experiments with polythene in various solvents [1] are in fair agreement with HUGGINS' and FLORY's theory. Discrepancies are attributed partly to the non-uniform molecular weight of the solute.

c. Physical meaning of interaction constant γ

A special merit of HUGGINS' and FLORY's theory is its general character. All specific aspects must therefore be concerned with the constant γ. From a purely theoretical point of view, this constant is determined by the interaction energies (see p. 63, $\gamma = \beta/kT$). In practice, however, several other factors play a role, and a number of shortcomings of the theory may be disguised by a proper choice of γ. For example, the theory neglected the effect of the interaction energy on the randomness of mixing. As shown by HUGGINS[2], this may be corrected for to a first approximation by a change in γ.

In fact, there is usually a perceptible discrepancy between the heat of mixing measured directly in the calorimeter, and that required by the value of γ [3]. Nor is γ in all cases rigorously independent of concentration (see Fig. 14A).

[1] R. B. RICHARDS, Trans. Faraday Soc., 42 (1946) 10.
[2] M. L. HUGGINS, Ann. N. Y. Acad. Sci., 44 (1943) 431.
[3] J. FERRY, G. GEE and L. R. G. TRELOAR, Trans. Faraday Soc., 41 (1945) 340.

As regards the dependence of γ on temperature, HUGGINS has shown that the following relation applies in a number of cases:

$$\gamma = A + B/T \qquad (71)$$

where A and B are constants. This is borne out by Fig. 14. To conclude, we quote some γ-values in Table 2.
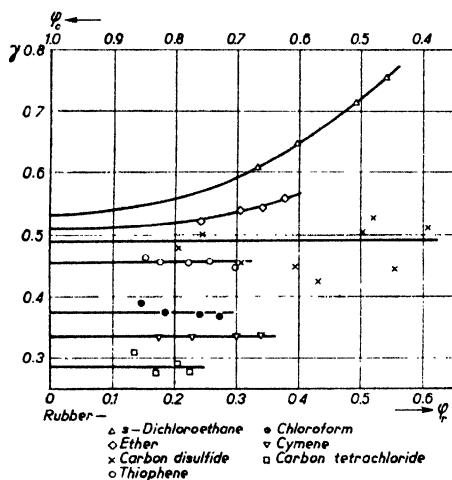
Fig. 14a.   Change of γ with concentration of rubber in various solvents.
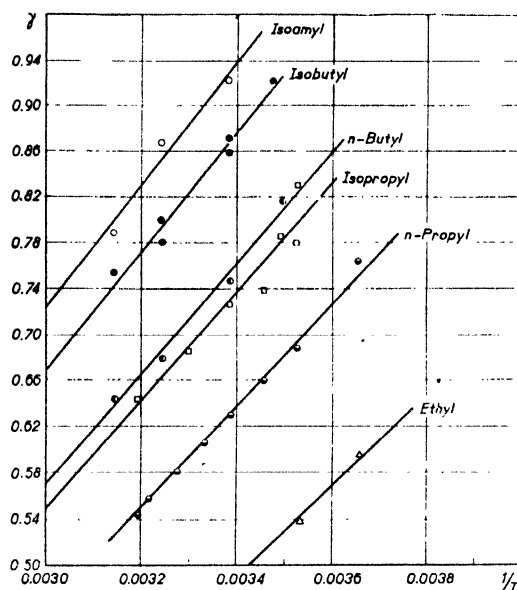
Fig. 14b.   Variation of γ with temperature for polystyrene in various laurates.

TABLE 2

VALUES OF γ IN VARIOUS MIXTURES

COMPONENTS	temp. ° C	γ
rubber-benzene + 10 % ethanol	25	0.26
,, -carbon tetrachloride	15—20	0.28
,, -cyclohexane	6	0.33
,, -chloroform	15—20	0.37
,, -light petroleum	25	0.43
,, -toluene	27	0.43
,, -benzene	25	0.44
,, -carbon disulphide	25	0.49
,, -benzene + 15% methanol	25	0.50
,, -ether	15—20	0.51
,, -symm. dichloroethane	15—20	0.53
polystyrene-benzene	5	0.2
,, -toluene	27	0.44
,, -n-butyl laurate	25	0.74
,, -isoamyl laurate	25	0.91
cellulose nitrate-cyclohexanone	25	0.15
,, ,, -acetone	± 25	0.2—0.3
cellulose acetate-tetrachloroethane	24.4	— 1.8

§ 8. OSMOTIC PRESSURE AND SWELLING PRESSURE

a. Osmotic pressure in dilute solutions

In dilute solutions of macromolecules the vapour pressure effect is usually too small to be of great practical interest. The osmotic pressure is then the only means of studying the thermodynamic properties of the solution. Since VAN 'T HOFF's law applies in the limit of infinite dilution, the osmotic pressure may be used to determine the molecular weight of the solute (compare chapter on molecular weight determinations, p. 133). In solutions of macromolecules, however, deviations from VAN 'T HOFF's equation are soon important even at low concentrations, and attempts have been made to derive expressions for the osmotic pressure at concentrations which are more easily accessible to experiment.

a. 1. Kinetic theories

Considerations on a purely kinetic basis were given by HALLER [1] as early as 1929. This author tried to extend the kinetic interpretation of VAN 'T HOFF's law to macromolecules, and assumed that the oscillations and rotations which the monomeric groups perform with respect to each other give rise to an additional osmotic pressure. On highly speculative grounds it was believed that these oscillations and rotations are only effective when two macromolecules cooperatively collide with the semi-permeable wall, so that the additional osmotic pressure is proportional to the square of the concentration and plays no part in the limiting law at great dilution.

OSTWALD [2] suggested a similar correction by writing

$$\pi = \frac{RT}{M}c + Kc^{\lambda} \tag{72}$$

where K and λ are constants, adding that usually λ is about 2. OSTWALD considered the additional term as representing a swelling pressure, because a formula of the type $\pi = kc^n$ was used by FREUNDLICH and POSNJAK [3] in earlier work to represent the swelling pressure in a number of gels. It is not clear, however, how a swelling pressure could be effective *in addition* to the osmotic pressure, and OSTWALD's equation must be considered chiefly as an empirical relation. This also applies to SCHULTZ's [4] formula. This author attempts to account for a volume correction which is analogous to VAN DER WAALS b-correction in the theory of gases: $p = RT(V-b)^{-1}$. In VAN 'T HOFF's law [5] V is replaced by M/c, which gives (if $b/M = s$),

$$\pi = \frac{RT}{M}\frac{c}{1 - sc} \tag{73}$$

In a number of solutions of globular proteins [6] this equation holds good with a constant value of s. The experiments show, however, that with flexible long-chain molecules the equation can only be maintained if s is assumed to be dependent on c. Now, since s represents the „effective volume" per gram of solute, SCHULZ assumes that FREUNDLICH and POSNJAK's equation holds good:

$$\pi = K s^{-\lambda} \tag{74}$$

This implies that the osmotic pressure π acts as a swelling pressure with regard to the effective volume s. *The constants K and λ are said to be independent of molecular weight M.* It is claimed by SCHULZ that his method allows of a more reliable extrapolation of osmotic data to zero concentration.

[1] W. HALLER, Kolloid-Z., 49 (1929) 74; 56 (1931) 257.

[2] Wo. OSTWALD, Kolloid-Z., 49 (1929) 60; Z. physik. Chem., 159 (1932) 375.

[3] H. FREUNDLICH and E. POSNJAK, Kolloid-Beih., 3 (1912) 442.

[4] G. V. SCHULZ, J. prakt. Chemie, 161 (1942) 147.

[5] The fact that VAN 'T HOFF's law is used as a starting point has sometimes given rise to the remark that this law is itself only an approximation to the logarithmic formula (27) on p. 59. In solutions of macromolecules, however, this is of no consequence, because the mole fraction x will always be small if the molecular weight of the solute is high, unless we are dealing with highly concentrated systems.

[6] H. H. WEBER and R. STÖVER, Biochem. Z., 259 (1933) 269.

Serious objections, however, were raised by several authors, in particular by DUCLAUX [1]. To begin with, equation (74) requires that the specific volume s becomes infinite in dilute solutions, with the result that the slope of the π/c versus c curve also becomes infinite (compare Fig. 15). Now, a possible check on the theory is to calculate s from (73) at a series of c-values and to plot log s against
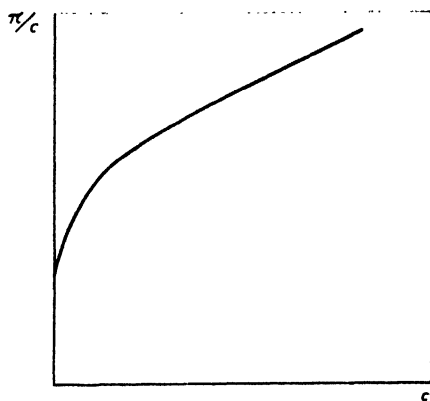
Fig. 15. Slope of π/c-c-curve at great dilution according to SCHULZ' theory.
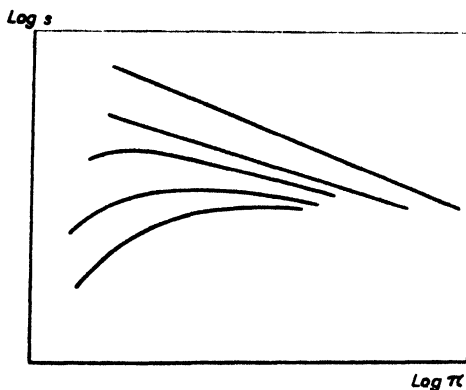
Fig. 16. Values of log s plotted against log π for various values of RT/M (cellulose nitrate in acetone).

the corresponding values of log π. This was carried out by DUCLAUX for cellulose nitrate in acetone with the result shown in Fig. 16. It is obvious that SCHULZ's method does no more than to determine an upper limit to RT/M. For all values of RT/M below 0.3 the curve in Fig. 16 is effectively linear. SCHULZ takes the highest RT/M-value, and motivates this by observing that lower values give rise to a π/c versus c curve which shows an "abnormal" decline at low values of c. This is clearly inconsistent, because this decline of the π/c-curve is an essential and immediate result of his own theory! Further, instead of equations (73) and (74) we may also write

$$\frac{\pi}{c} = \frac{RT}{M} + K^{1/\lambda} \pi^{1 - 1/\lambda}$$

As K and λ are said to be independent of M, a plot of π/c against π in a homologous series should give parallel curves. This is by no means borne out by experiment.

DUCLAUX himself [2] gives a kinetic picture of the osmotic pressure in polymer solutions, which is of some interest as an interpretation of the entropy effects discussed in the preceding sections. He observes that VAN 'T HOFF's law holds good when the probability for the presence of a molecule at a certain point is not affected by the presence of neighbouring molecules. This condition is violated in the case of anisodiametric particles, because they hinder each other's Brownian movement. This effect is independent of cohesive forces between the particles, nor has it anything to do with the co-volume, because it would be equally effective with infinitely thin threads. In fact, it does no more than to give a simple qualitative geometric interpretation of the effect which a polymer molecule has on the possible positions of neighbouring molecules. (See entropy considerations on p. 73.)

a. 2. Thermodynamic treatment

The thermodynamic formula for the osmotic pressure (p. 59) is

$$\pi = \frac{-\Delta g_o}{v_o}$$

where v_o is the molar volume of the solvent, and $\Delta g_o = g_o - g_o^\circ$ is the difference

---

[1] J. DUCLAUX, *J. chimie physique*, 41 (1944) 209.
[2] See also A. POLSON, *Nature*, 157 (1946) 406.

between the molar free energy of the solvent in the mixture and that of the pure solvent. From equation (61) it follows that

$$\pi = -\frac{RT}{v_o}[\ln q_o + (1 - \frac{1}{n})q_r + \gamma q^2_r]$$ (75)

As has already been observed on p. 74, when q_r is small compared with unity, we may omit powers of q_r higher than the second. This gives

$$\pi = \frac{RT}{v_o}[\frac{1}{n}q_r + (\frac{1}{2} - \gamma)q^2_r].$$

Since there are N_r solute molecules with molecular weight M in a volume $(N_o + nN_r)v_o$, the concentration in grams per unit of volume is $c = Mq_r/v_o n$. Consequently

$$\pi = \frac{RT}{M}c + RTv_o\left(\frac{n}{M}\right)^2(\frac{1}{2} - \gamma)c^2$$ (76)

This shows that a plot of π/c against c should give a straight line. For a given polymer type in a given solvent, *the slope of this line is independent of the molecular weight*, since γ is independent of M and n is proportional to M. Examples are given in Fig. 17.
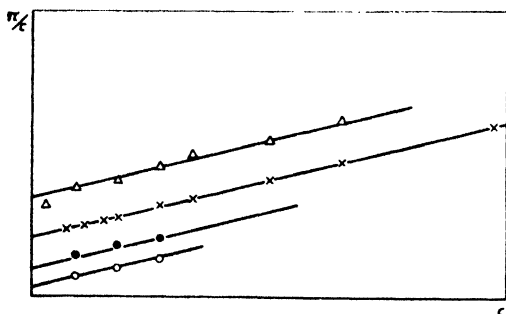
Fig. 17a. Reduced osmotic pressure π/c as a function of c in acetone solutions of nitrocellulose [1].
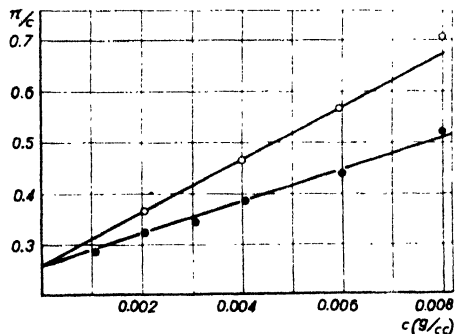
Fig. 17b. Reduced osmotic pressure π/c for gutta-percha solutions in toluene (●) and in carbon tetrachloride (o) [2]

In some cases it is essential to add a third order term, since this does perceptibly affect the extrapolated value of π/c at zero concentration[3]. In most cases, however, the equation (76) will suffice, giving a theoretical foundation to the linear extrapolation of π/c versus c. Deviations from the straight line will, of course, become apparent when φ_r is no longer small compared with unity. A discussion of more concentrated polymer solutions, however, is given in section 8 b.

It is interesting to observe that the second term in equation (76) vanishes when $\gamma = \frac{1}{2}$. In that case VAN 'T HOFF's law holds good in a considerable range of concentrations. Such solutions are, however, properly regarded as being *less* and not *more* "ideal" than those in which $\gamma = 0$.

[1] G. V. SCHULZ, Z. physik. Chem., A 176 (1936) 317.
[2] H. STAUDINGER and K. FISCHER, J. prakt. Chem., 157 (1940) 19, compare M. L. HUGGINS, Ind. Eng. Chemistry, 35 (1943) 216.
[3] Compare M. L. HUGGINS, Ind. Eng. Chem., 35 (1943) 216.

b. Osmotic or swelling pressure in concentrated systems

b. 1. General considerations

At high concentrations most polymer solutions assume gel-like behaviour. (Compare chapter on gels, p. 494). The osmotic pressure of a gel is usually called swelling pressure, although from a physical point of view these quantities are identical: they both represent the pressure difference between solution and solvent separated by a semipermeable membrane when equilibrium is reached. In particular, the vapour pressure p and the swelling pressure π of a gel are related to the free energy of dilution according to the usual formula

$$RT \ln \frac{p}{p_o} = \varDelta g_o = - v_o \pi,$$

where p_o is the vapour pressure of the pure solvent and v_o the molar volume of the solvent in the mixture, (properly speaking, the average value of the molar volume for pressures ranging from that in the pure solvent to that in the solution). From the value of $\varDelta g_o$ one may derive the heat of dilution

Fig. 18. Osmotic pressure of polymer solution in dependence on volume fraction φ_r for various values of the interaction constant γ.

$$\varDelta h_o = \frac{\partial(\varDelta g_o / T)}{\partial(1/T)},$$

e.g., by plotting $\varDelta g_o / T$ against $1/T$. This heat of dilution in gels is usually called *differential heat of swelling*, in contrast to the *integral* heat of swelling $\varDelta H$ which represents the total heat absorbed on mixing the dry gel with the solvent. The values of $\varDelta H$ and $\varDelta h_o$ may, of course, also be obtained from direct calorimetric determinations. According to KATZ[1], who made some very fundamental researches on the thermodynamics of gels, the value of $\varDelta H$ is practically equal to that of $\varDelta G$, which means that the entropy of mixing in gels is negligible compared with the energy effects. In such cases one speaks of "ideally concentrated" systems (NERNST). Later investigations have shown that $\varDelta S$ is only negligible in special cases where $\varDelta H$ is negative and large. What is more, swelling may take place notwithstanding a positive heat of swelling, i. e., if heat is absorbed in the swelling process. This is apparent in particular from GEE's investigations in the physical chemistry of rubber. This author showed that the

---

[1] J. R. KATZ, Ergeb. exakt. Naturw., 3 (1924) 316; 4 (1925) 154.

swelling of raw rubber is best considered as a problem of miscibility and may be treated by the methods of HUGGINS and FLORY.

In this connection, it is interesting to consider the osmotic pressure of polymer solutions at high concentrations[1]. This is especially illustrative of the similarity between FLORY's theory of polymer solutions on the one hand and VAN DER WAALS theory of gases on the other hand. Fig. 18 shows how the osmotic pressure changes with concentration for various values of γ. For convenience the low number 10 was chosen for the number of segments in the polymer molecule, but curves of a similar nature are obtained with higher values of n. For a given critical concentration of γ, ($\gamma = \gamma_c$), the curve shows an inflexion point with horizontal tangent. If $\gamma > \gamma_c$, the curves show an unstable region corresponding to the miscibility gap of Fig. 11. The analogy with VAN DER WAALS' isotherms is obvious.

Interesting experiments on the equilibrium swelling of rubber are recorded by GEE[2]. We know from the preceding sections that the entropy of dilution in polymer solutions is independent of the solvent; it depends only on the degree of polymerisation and on the volume fraction of the solute. (See, for instance, equation 56, p. 74). This means that the swelling power of a liquid for a given polymer will be determined by the value of the interaction constant γ. In view of equation (34), which relates $\beta(= \gamma\ kT)$ to the density of the cohesive energy in the solvent, we must expect that the swelling of rubber in various liquids is governed by this cohesive energy density. Fig. 19 shows a plot of equilibrium swelling against cohesive energy density according to GEE[2]. The maximum swelling is attained when the cohesive energy density of the rubber is the same as that of the solvent. A more detailed discussion is given in the papers quoted.
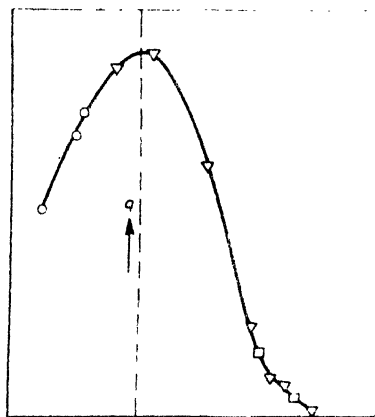
Fig. 19. Equilibrium degree of swelling q of pure gum rubber in various liquids plotted against cohesive energy density of the liquid.

○ hydrocarbons, ▽ esters, ☐ ketones.

b. 2. Cross-linked gels

So far we have tacitly assumed that the polymer substance is free to swell at liberty. In many gels, however, the polymer molecules are interlinked in junction points which are not accessible to the swelling agent. These links may represent primary valence bonds (e.g., sulphur bridges in vulcanised rubber), or secondary valence bonds; the characteristic feature of the gel in both cases is the *network structure*, which impedes unlimited swelling. The influence of this reticular structure on the thermodynamic properties of a gel may be accounted for by adding the "free energy Φ of the network" to that of the mixture as such. On swelling, this free energy of the network is increased, which accounts for a smaller swelling tendency. This may also be expressed by saying that the osmotic pressure of the gel is decreased by an amount $\partial\Phi/\partial V$, where V is the volume of the gel.

[1] P. DEBYE, *Report Cornell University.*

[2] G. GEE, *Trans. Faraday Soc.*, 38 (1942) 418; *Trans. I.R.I.*, 18 (1943) 266.

The molecular origin of Φ may be seen as follows. When the volume of the gel is increased, the structural units of the gel scaffolding connecting successive junction points are brought into configurations of lower entropy (and possibly also of higher energy). They tend to reassume the original configuration to attain minimum GIBBS free energy. For special models, Φ may be calculated from molecular structure. An example is given in the theory of random coiling of long chain molecules (see p. 127), where it is assumed that changes in Φ are due to changes in the configurational entropy of the network (energy effects being zero).

The situation may be further explained as follows. We consider a gel of a polymer substance free from interlinking. In principle, therefore, this gel may show unlimited swelling (in a suitable solvent, that is). Suppose now that this gel contains m moles of solvent per gram of polymer and is in equilibrium with the solvent vapour through a semipermeable membrane (see Fig. 20). Let p be the pressure in the vapour phase and σ that in the gel. The equilibrium requires that the partial GIBBS free energies of the solvent are equal in the two phases: $g_o = g'_o$. This represents a relation between p, σ and m. Changing these quantities at constant temperature we obtain

Fig. 20. Equilibrium between a gel at hydrostatic pressure σ and solvent vapour at pressure p through a semipermeable membrane.

$$v_o \, d\sigma + \left(\frac{\partial g_o}{\partial m} \right)_p dm = v_p \, dp \qquad (77)$$

where $v_o = (\partial V / \partial m)_p = (\partial g_o / \partial \sigma)_m$ is the molar volume of the solvent in the gel and v_p is the volume per mole of solvent in the vapour, i.e., approximately, RT/p. The equation (77) determines the change in vapour pressure as a result of changes in hydrostatic pressure σ in the gel, or of changes in the amount of solvent absorbed. In particular we have, for $dm = 0$,

$$\frac{RT}{p} \left(\frac{\partial p}{\partial \sigma} \right)_m = v_o \; .$$

This equation may be used to reduce the vapour pressure of a gel to zero hydrostatic pressure in the gel:

$$RT \, \ln \frac{p}{p_o} = v_o \, \sigma \qquad (78)$$

Suppose now, that we are dealing with an interlinked gel in equilibrium with solvent vapour. The constraints imposed upon the scaffolding structure when the gel swells to a greater volume have the same effect as a hydrostatic pressure. One way of accounting for them has been mentioned already: we include the GIBBS free energy Φ of the network structure in the GIBBS free energy of the whole system; the equilibrium swelling will then be determined by the minimum value of this total GIBBS free energy:

$$\partial G / \partial m = 0.$$

This reasoning was followed by FLORY and REHNER[1] in their analysis of swelling networks. We may also calculate the equilibrium from a free energy which does not contain Φ, provided the result is corrected afterwards by means of equation (78). Clearly, this method is equivalent, since $\partial \Phi / \partial m = (\partial \Phi / \partial V) v_o = - \sigma v_o$, where σ represents the hydrostatic pressure due to network interlinking. BARKAS[2] and CASSIE[3] have used this method to correct the sorption isotherm of wood and of wool. In Fig. 21 the drawn curve represents the sorption of water by wool measured directly.

The swelling involved in this sorption, however, is opposed by forces whose magnitude may be estimated from stress-strain data in the stretch of wool. If the vapour pressure is corrected according to equation (78), one finds the much lower vapour pressures of the dotted curve. This curve extends no further than to 50% relative vapour pressure. Physically this means that the water in the fully swollen wool tends to be "squeezed" out of the gel by the forces acting in the gel structure, and is, for this reason, in equilibrium with a vapour of relative vapour pressure 100% instead of 50%. Incidentally we observe that the sigmoid shape of the isotherm does not persist in the corrected curve. It is to be noted, however, that the application of stress-strain data to the swelling process is of an extremely approximate nature, and the results recorded here are not to be considered as final.
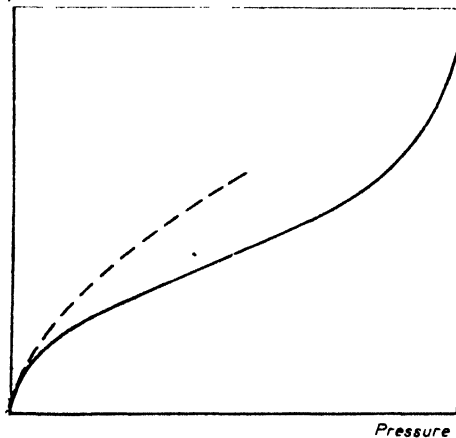
Fig. 21. Vapour pressure of wool at various regains, —— observed, - - - - corrected for hydrostatic pressure in the gel resulting from mechanical contraints in the structure, which oppose the swelling[3].

The behaviour of cross-linked gels in swelling media has been studied experimentally by MAGAT, by BOYER[4] and by GEE[5]. In particular it is interesting to note the following application. If a swollen gel in equilibrium with pure solvent is transferred to a solution of a high polymer in the same solvent, the gel will deswell until a new equilibrium is established. The extent of deswelling depends on the degree of cross-linking in the gel and on the activity of the solvent in the polymer solution. This activity is determined by the amount and the molecular weight of polymer dissolved and on the interaction constant γ. Conversely, from the extent of shrinking observed when the gel is brought into contact with the polymer solution one may derive the molecular weight of the polymer dissolved. The special merit of this method, which admittedly is not very accurate, lies in its simplicity.

[1] P. J. FLORY and J. REHNER, *J. Chem. Phys.*, 11 (1943) 521.

[2] W. W. BARKAS, *Forest Product Research*, no. 6 (1945); *Trans. Faraday Soc.*, 38 (1942) 194

[3] A. B. D. CASSIE, *Trans. Faraday Soc.*, 41 (1945) 458.

[4] R. F. BOYER, *J. Chem. Phys.*, 13 (1945) 363.

[5] G. GEE, *Trans. Faraday Soc.*, 42 B (1946) 33.

§ 9. REVIEW OF THE THERMODYNAMICS OF POLYMER SOLUTIONS

a. Limits of applicability

The rather general character of the thermodynamic theory of polymer solutions makes it particularly suitable to a variety of substances. It gives a satisfactory explanation of the most characteristic properties of these solutions. In particular, it describes in a satisfactory manner the phenomena of separation into two layers, one of which is very often a gel. This phenomenon is the starting point in the theory of *coacervation*, which plays such an important part in biology (compare the chapter on colloids with electrolyte character, p. 184), and it is tempting to apply the thermodynamical principles developed to these systems. We have reason to believe that this application is permissible in principle, but caution is required when results of an accurate quantitative nature are wanted. The theory given so far makes use of a number of approximations which are likely to fail in systems containing electrolytes. The only quantity which is characteristic for the system is the interaction constant γ, whose role in the phenomenon of separation is decisive. It is obvious that the very specific interaction of strongly polar or ionogenic groups cannot be described satisfactorily by a VAN LAAR heat of mixing. In particular the assumption that mixing is approximately random becomes very unsatisfactory. Discrepancies are already observed in such systems as polythene-nitrobenzene, where the nitrogroup of the solvent is responsible for deviations from theory [1]. The precipitation of proteins by salts, for instance, can be explained qualitatively by the theory of polymer precipitation (p. 144), but we should be on our guard against too specified a quantitative statement. The very specific influence which the solvent may have on the properties of proteins is well shown by SVEDBERG's experiments on the relation between pH and molecular size [2]. It is obvious that such pronounced effects cannot be described by a simple VAN LAAR interaction term.

Further discrepancies may result from effects such as those mentioned on p. 64 and embodied in LANGMUIR's equation (37). Besides, in this chapter we have gone no further than to give a brief outline of the thermodynamic theory. More detailed accounts and extensions to more special cases can be found elsewhere [3]. In particular we mention GUGGENHEIM's [4] extension to mixtures in which the solvent molecules also occupy more than one lattice point.

Critical considerations of HUGGIN's and FLORY's approach were given by FLORY [5] and in particular by ZIMM [6]. The latter author applied the statistical method developed by MAYER [7] and once more stressed the important fact that the experimental constant γ does not merely account for the interaction energy but also serves to mask the shortcomings of the present theory.

---

[1] R. B. RICHARDS, *Trans. Faraday Soc.*, 42 (1946) 10, 20.

[2] See also the chapter on colloids with electrolyte character.

[3] E. A. GUGGENHEIM, *Proc. Roy. Soc.*, A 183 (1944) 203; *Trans. Faraday Soc.*, 41 (1945) 107; A. R. MILLER, *Proc. Cambr. Phil. Soc.*, 38 (1942) 109; 39 (1943) 54.

[4] E. A. GUGGENHEIM, *Nature*, 153 (1944) 406.

[5] P. J. FLORY, *J. Chem. Phys.*, 13 (1945) 453.

[6] B. H. ZIMM, *J. Chem. Phys.*, 14 1946 164.

[7] J. E. MAYER, *J. Chem. Phys.*, 9 (1941) 2; 13 (1945) 276.

b. Discussion of some experiments

It is not our aim to give a detailed discussion of experimental results obtained in the thermodynamical study of polymer solutions. We will do no more than consider a few examples. Among the earlier measurements we may mention the calorimetric determinations of the heat of dilution in solutions of cellulose nitrate in various solvents [1] and the heat of solution of cellulose acetate in methyl acetate [2]. In both these cases heat was generated on mixing ($\triangle H < 0$). Later studies in the system cellulose nitrate-acetone [3,4] have shown that the heat of dilution is positive at high concentrations of cellulose nitrate and becomes negative at great dilution. The *concentrated* system shows thermodynamic properties which are practically independent of the molecular weight M of the cellulose nitrate, in conformity with theory. This is obvious from Fig. 22, which gives the free energy $\varDelta g_0$ of dilution plotted against percentage by weight. At about 55% the slope of the $\varDelta g_0$ curve shows a sudden change. This indicates the occurrence of a polymer-solvent compound with 6 molecules acetone per glucose-unit, in conformity with X-ray investigations [4]. It is to be noted in this connection, that the entropy of mixing in the system cellulose nitrate-acetone becomes *negative* in the limit of very low concentrations of acetone. This is explained by SCHULZ on the assumption that the solvent molecules are absorbed by the cellulose nitrate on to localised sites (perhaps in an orientated state), to which one would have to attribute a lower entropy. A behaviour of this type is, of course, not accounted for in HUGGINS' and FLORY's theory of the entropy of mixing.

As regards osmotic pressure, we may refer to p. 85, in particular to Fig. 17. The fact that the slope of π/c versus c is independent of M is in good

Fig. 22. Free energy of dilution per mole acetone in the system cellulose nitrate-acetone plotted on a logarithmic scale against percentage by weight of the cellulose nitrate.

agreement with theory and is confirmed by a number of further investigations [5]. This fact was observed by various other authors, among whom we mention KRATKY and MUSIL [6], and is of particular importance to the method of extrapolation to zero concentration. SCHULZ' extrapolation method was discussed on p. 83; its physical significance is obscure and its practical applicability doubtful [7].

[1] I. OKAMURA, *Kolloid-Z.*, 65 (1933) 175.
[2] LIEPATOW and PREOBRAGENSKAJA, *Kolloid-Z.*, 68 (1934) 324.
[3] G. V. SCHULZ, *Z. physik. Chem.*, A 180 (1937) 1; 184 (1939) 1; B 40 (1938) 319; 52 (1942) 253; *Z. Elektrochem.*, 45 (1939) 652;.
[4] E. CALVET, *Compt. rend.*, 212 (1941) 542; 213 (1941) 126; 214 (1942) 716, 767.
[5] M. L. HUGGINS, *J. Am. Chem. Soc.*, 64 (1942) 1712; *Ind. Eng. Chem.*, 35 (1943) 216.
[6] O. KRATKY and A. MUSIL, *Z. Elektrochem*, 43 (1937) 326.
[7] See, for example, G. GEE, *Trans. Faraday Soc.*, 36 (1940) 1171.

Extensive studies in the physical chemistry of rubber solutions were carried out by GEE[1]. We may mention further RICHARDS' investigations on polythene solutions[2]. GEE's experiments were of particular importance because he was the first to show that high polymer substances may be dissolved with absorption of heat, contrary to the current opinion[3]. The results of these experiments are in good agreement with theory; see also the text of the preceding sections.

A further discussion of some experiments is given in the chapter on gels (p. 558). Recent investigations have also made use of light scattering to obtain thermodynamic data. This method is referred to in the chapter on molecular weight determinations (p. 146) in which we also consider the osmotic method (p. 133). The experiments on coacervation are discussed in great detail in Chapter VIII, p. 232.

[1] G. GEE and L. R. G. TRELOAR, *Trans. Faraday Soc.*, 38 (1942) 147; G. GEE, *ibid.*, p. 276 and p. 418; G. GEE, *Trans. Faraday Soc.*, 40 (1944) 463; *Transactions I.R.I.*, 18 (1943) 266; J. FERRY, G. GEE and L. R. G. TRELOAR, *Trans. Faraday Soc.*, 41 (1945) 340. G. GEE and W. J. C. ORR, *Trans. Faraday Soc.*, 42 (1946) 507.

[2] R. B. RICHARDS, *Trans. Faraday Soc.*, 42 (1946) 10, 20.

[3] See, for example, F. EIRICH and H. MARK, *Erg. Ex. Naturw.*, 15 (1936) 1; G. V. SCHULZ, *loc. cit.*

IV. THE PHYSICAL PROPERTIES OF RANDOMLY KINKED LONG CHAIN MOLECULES

J. J. HERMANS

Laboratory for inorganic and physical chemistry, University of Groningen.

§ 1. INTRODUCTION

This chapter will be concerned with a number of physical properties which are closely connected with the shape of long chain molecules in solution and the solid state. For simplicity we shall consider chains which are composed of a certain (large) number of repeating units, the monomer groups. Let the molecular weight of a monomer group be M_g. It would of course be possible to extend our considerations to those cases where more than one kind of monomer unit occurs in the chain, in which case M_g represents an average value. This would however greatly complicate the treatment without giving an important gain in its generality. We shall define the degree of polymerisation P as the ratio between the molecular weight M of the macromolecule and that of the monomeric group. Thus

$$M = P \ M_g \tag{1}$$

As a simple example of a linear macromolecule let us consider a paraffin chain. X-ray investigations have shown that the molecular chains in solid paraffins or fatty acids are lying in the stretched zigzag arrangement (Fig. 1a). This arrangement is particulary favourable on energetic grounds on account of the interaction with neighbouring molecules. It is clear, however, that a great many shapes are possible while retaining the valency angle of 109°. In Figure 1b such a possible shape has

Fig. 1a. Paraffin chain, stretched.   Fig. 1b. Paraffin chain, twisted.
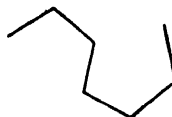
been reproduced. The reader may bear in mind that a great many more arrangements are possible in three dimensions than can be drawn in two. On the strength of this argument we must expect that chain molecules, when moving freely in a liquid, will occur in a large number of different forms and that, as a result of temperature movement, they will constantly be changing from one shape into another. We may express this by saying that the macromolecule behaves like a loosely built flexible "skein" or "coil".

Not all arrangements observant of the valency angle are equivalent from an energetic point of view. Even in the simple case of paraffin chains, a certain steric

hindrance exists: the rotation round about a line connecting two successive carbon atoms is not completely free. Out of all possible molecular shapes, however, we may select those of equal energy. If the degree of polymerisation is sufficiently high, the number of these arrangements is still very large. Generally two of such arrangements will be separated by a potential barrier, and this has given rise to serious objections to the conceptions introduced here [1]. By way of illustration let us consider a simple molecule such as ethane. If in this molecule one of the methyl groups is rotated with respect to the other round the line joining the C-atoms, maxima and minima in the potential energy are encountered, which alternate every 60°. An extreme value of the energy is acquired when the hydrogen atoms of the one group are exactly opposite those of the other. It is not yet known whether this represents the state of maximum or of minimum energy. According to earlier calculations of EYRING [2] the height of the potential barrier between two successive minimum values is about 400 cal/mol, which tallies well with the measurements of EUCKEN [3], who derived a value of 315 cal/mole from the behaviour of the specific heat at low temperatures. If this were true, the potential barrier would be no serious obstacle to rotation, since the thermal energy at ordinary temperatures is also of the order of 300 cal/mole. Later measurements [4], however, have revealed that the potential barrier is of the order of 3 000 rather than 300 cal/mol. The rotation is, therefore, restricted mainly to oscillations through a fairly small angle, and among, say, 1 000 oscillations there is perhaps but one which results in a transition [4]. Yet, even in this case the number of transitions per second is very large, since the frequency of the oscillation is extremely high. Thus we may expect that flexible chains occur even in those cases where considerable potential barriers are present: it must be remembered that a time of the order of, say, 10^{-8} sec is a very long time from the point of view of molecular processes.

In recent papers KUHN has considered the influence of the presence of potential barriers on the time which a long chain molecule needs to change its shape to a considerable extent. These considerations lead to the concept "inner viscosity" of the molecular chain, which determines the velocity with which changes in shape will take place under the influence of external forces. A brief account is given on p. 114 in connection with the treatment of the viscosity contribution of chain molecules in solution.

§ 2. STATISTICAL TREATMENT. OPTICAL ANISOTROPY OF THE MOLECULE

In the following sections we shall consider the isolated macromolecule in solution. In other words, the concentration of the solute should be so small, that each molecule with its immediate surroundings may be treated as an isolated system. Some properties

[1] S. BRESLER and J. FRENKEL, Acta Physicochim. URSS, 11 (1939) 485. G. BIER, Experientia, 2 (1946) 82.

[2] H. EYRING, J. Am. Chem. Soc., 54 (1932) 3191.

[3] A. EUCKEN et al. Z. physik. Chem., B 20 (1933) 184; 23 (1933) 265.

[4] R. K. WITT and J. D. KEMP, J. Am. Chem. Soc., 59 (1937) 273; W. HUNSMANN, Z. physik. Chem. B, 39 (1938) 23; KISTIAKOWSKY, LACHER, and STITT, J. Chem. Phys., 6 (1938) 407. Compare the calculations of K. SCHÄFER, Z. physik. Chem., B 40 (1938) 357, and the estimates by E. GORIN, J. WALTER and H. EYRING, J. Am. Chem. Soc., 61 (1939) 1876. It is to be noted that double bonds enhance the free rotation round neighbouring single bonds: M. L. HUGGINS, J. Polym. Sci., 1 (1946) 1.

of concentrated solutions will be the subject of Section 9; those of the solid substance will be treated in Section 10.

a. Introductory remarks

Let us now try to describe in more detail the shape and flexibility of the molecules. Usually the nature and magnitude of the factors restricting free rotation are unknown or known incompletely. In order to arrive at a general theory, KUHN[1] has therefore introduced the so-called statistical chain-element, which will be called chain element for short.

Imagine a monomeric group (AB in Fig. 2) fixed in space. Then the orientation of the next monomeric group, although not determined completely, will be restricted to a relatively small region of space angles. In the simple case of completely free rotation the atom C may lie anywhere on the circle CC'. For the atom D, next to C, a large number of positions is available, and it is obviously possible to indicate a certain number of successive monomeric groups in such a manner that the orientation of the last of these groups shows practically no correlation with the orientation of the first. The smallest possible number of successive monomeric groups with this property forms a statistical chain-element. We will designate this number by ν, i.e., we imagine the molecule to be split up into

$$N = P/\nu \qquad (2)$$
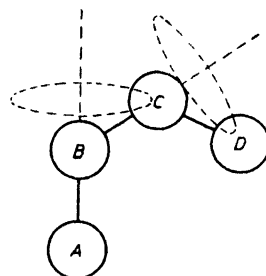
Fig. 2. Possible positions of successive atoms.

chain-elements, each containing ν monomeric groups. The orientation of a chain-element in space is practically independent of the orientation of all other chain-elements.

It cannot be denied that the introduction of these chain-elements is, to a certain extent, arbitrary. They enable us, however, to treat a number of widely differing macromolecules on a common basis. Moreover, the value of ν can be estimated from experimental data (compare the ensuing sections). And not only is approximately the same value derived from a number of independent experiments, but also this value is quite acceptable and it correlates with the properties of the macromolecules concerned in an expected manner. One should, for example, expect that the number ν may be chosen much smaller in saturated paraffins than, for instance, in cellulose derivatives. In the latter the free rotation is impeded to a much greater extent by steric hindrance and by the interaction of hydroxyl groups. For similar reasons the number ν in cellulose nitrate becomes larger as more NO_2 groups are substituted in the glucose residue (see below).

Recently HUGGINS[2] has pointed out that the flexibility of macromolecules may also be influenced by the grouping of side chains along the main chain. For instance, if the monomeric units are of the type CHR, where R represents a substituent, there are two different, non-equivalent dispositions of H and R relative to the plane CCC in Fig. 3. This disposition cannot be changed by mere rotation around single bonds. This suggestion
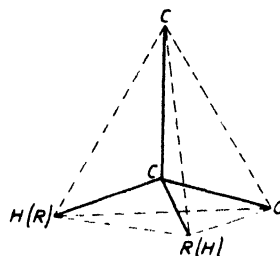
Fig. 3. Two different positions of H and R along the main chain -CCC-

[1] W. KUHN, Kolloid-Z., 68 (1934) 2; 76 (1936) 258.
[2] M. L. HUGGINS, J. Am. Chem. Soc., 66 (1944) 1991.

is of interest in that it could explain the influence of the temperature of polymerisation on a number of properties, without resorting to the assumption of ramification. The higher the temperature in the polymerisation process, the greater is the tendency of the side chains to show a completely random sequence.

Let us assume that to each chain-element we may attribute a length A. Strictly speaking, A is only an average length, since the monomeric groups in a chain-element are not always arranged in the same manner. To a first approximation, however, we may assume that all chain-elements are of equal length. Similarly, let us attribute to each chain-element two principal polarisabilities a_1 and a_2, which as a rule will be different: the chain-element is optically anisotropic. We call a_1 the polarisibility in the direction along the chain, a_2 the one perpendicular to this direction. As regards the anisotropy $a_1 - a_2$, we know only its order of magnitude. It will be smaller than v times the anisotropy of the monomeric group, since the monomeric groups in a chain-element on the average are arranged in a somewhat irregular kinky manner. Nevertheless, the experimental value of $a_1 - a_2$ will lead to a rough estimate of v if the anisotropy of the monomeric group is known (compare Section 10 of this chapter).

It is further to be expected that the value of $a_1 - a_2$ depends to a certain extent on the solvent molecules surrounding the chain-elements. In fact, the entropy change at swelling[1] indicates that some of the solvent molecules may be oriented by the solute molecules, which is likely to give rise to an extra birefringence ("birefringence of adsorption", according to VERMAAS[2]).

b. Statistical considerations

After these preparatory remarks let us consider more closely the spatial arrangement of the statistical chain-elements. They have the property that their orientations in space are mutually independent. According to each arrangement of successive chain-elements, there results a corresponding different total "length" r of the macro-molecule. By this length r is meant the distance between the two ends of the molecule, and we shall first consider all molecules of a given length r. The arrangement of the N chain-elements is, of course, not the same in all these molecules. All we can say is that the sum of the projections of all N chain-elements in the direction of r must equal the length of the molecule. Thus, if the orientation of the i th chain-element is characterised by the angle ϑ_i between this chain-element and the direction of r (Fig. 4), we have:
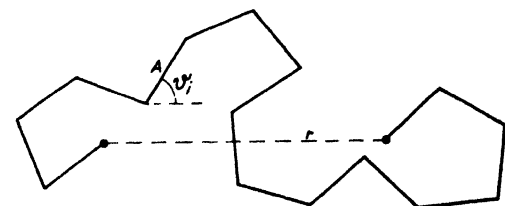
Fig. 4. Orientation of the chain-elements with respect to the vector r, joining the ends of the molecule

$$r = A \; \Sigma_i \cos \vartheta_i$$

We may now ask: what is *the most probable* distribution of the directions ϑ_i over the N chain-elements ? The answer to this question can easily be given by analogy with a similar problem in molecular

[1] W. HALLER, *Kolloid-Z.*, 56 (1931) 257. G. V. SCHULZ, *Z. physik. Chem.*, B 52 (1942) 253. See also Chapter III, p. 91)
[2] D. VERMAAS, *Z. physik. Chem.*, B. 52 (1942) 131. See also H. NEUBERT, *Kolloidchem. Beihefte*, 20 (1925) 244; A. MÖHRING, *Kolloidchem. Beihefte*, 23 (1927) 152.

statistics. To that end we need only recall how an energy E is distributed over a large number N of molecules. It is well-known, that the most probable distribution is that where the number of molecules with energy ε_i is proportional to

$$e^{\beta^1 \iota},$$

Here the factor β^1 is connected with the average energy E/N per molecule. It is shown in molecular statistics that

$$\beta^1 = - 1/kT$$

In the present case the situation is completely analogous. The quantity r/A has to be distributed over N chain-elements; the contribution of the i th chain-element to this quantity is $\cos \vartheta_i$. Accordingly, the probability of finding an angle ϑ_i is proportional to [1]

$$e^{\beta \cos \vartheta_i} \tag{3}$$

The factor β is again determined by the average contribution per chain-element, i.e., by r/NA.

Knowing the distribution of the chain-elements over the various directions in space, it will not be difficult to express the anisotropy of the macromolecule as a whole by that of the chain-element. We shall omit the calculation [1] and mention only the anisotropy $\gamma_1 - \gamma_2$ of the macromolecule in the special case when r/NA is small:

$$\gamma_1 - \gamma_2 = {}^3/_5 (\alpha_1 - \alpha_2) \frac{r^2}{NA^2} \tag{4}$$

Here γ_1 is the polarisability of the macromolecule in the direction of r and γ_2 that perpendicular to r. The result depends on the value of r, as the chain-elements are more and more oriented in the direction of r as r increases. It is clear that this anisotropy will not become apparent as long as the macromolecules are left to themselves. For, in a sample of many molecules with length r, the direction of the vector r is uniformly distributed over all directions in space. It is only when some orientations of r have preference above others that the anisotropy is measurable. (Compare Sections 7 and 10 of this chapter.)

So far we have considered only molecules of a given length r. In a sample of many molecules all possible values of r will occur, and we shall now study the frequency distribution as a function of r. The probability of finding a distance r between the two ends of the molecule will be proportional to the *number of arrangements* by which this distance can be achieved. It is obvious that the maximum distance NA can only be reached in one way, viz., by arranging all chain-elements one after another in a straight line. With decreasing value of r, the number of possible arrangements will increase, and it will be shown that this number is given approximately by a GAUSS-ian distribution curve. To that end let us consider a one-dimensional analogy [2]. Suppose we agree to make N successive steps of length A along a straight line. Whether a step is to be made in the positive or the negative direction is decided by a "head

---

[1] W. KUHN and F. GRÜN, *Kolloid-Z.*, 101 (1942) 248.

[2] E. GUTH and H. MARK, *Monatshefte für Chemie*, 65 (1935) 93; RAYLEIGH, *Scientific Papers*, Vol. VI, p. 600.

or tail" lottery. The distance reached at the end of the Nth step is A times the difference between the numbers of positive steps and negative steps respectively. The probability of reaching a distance mA is thus proportional to the number of ways in which among N steps, there are $\frac{1}{2}N + \frac{1}{2}m$ positive ones and $\frac{1}{2}N - \frac{1}{2}m$ negative ones, irrespective of the order of succession. This number is

$$\Omega = \frac{N!}{(\frac{1}{2}N + \frac{1}{2}m)! \ (\frac{1}{2}N - \frac{1}{2}m)!}$$

It has its maximum value at $m = 0$, and in the neighbourhood of this maximum approximates closely[1] the GAUSSEAN probability function:

$$\Omega = ae^{-m^2/2N}$$

The number a may be so chosen that a summation over all possible values of m gives unity as a result. The quantity Ω then represents the *probability* of finding a distance mA. Consequently, the probability of reaching a distance x is

$$ae^{\dfrac{-x^2}{2\,NA^2}} \tag{5}$$

In the long-chain molecule any of the N chain-elements is equivalent to a step in three-dimensional space, the direction of which is entirely arbitrary. The three-dimensional formula differs but slightly from the formula (5), and may be expressed as follows:[1, 2] the probability of finding one end of the molecule in a volume-element (dx dy dz) at distance r from the other end equals

$$\Psi_o \ dx \ dy \ dz = ae^{\dfrac{-3r^2}{2\,NA^2}} \ dx \ dy \ dz \tag{6}$$

It will be clear that the probability of finding a distance r, irrespective of the direction in space, becomes

$$W(r)dr = a.4\,\pi\,r^2 e^{\dfrac{-3r^2}{2\,NA^2}} . dr \tag{7}$$

The constant a is determined by the fact that the integral of (7) over all possible values[3] of r equals 1. In Fig. 5 the function $W(r)$ is shown for two values of N. It is seen that the top of the curve is shifted towards larger values of r if N is increased, but this shift is only proportional to $\sqrt{N}$. This also applies
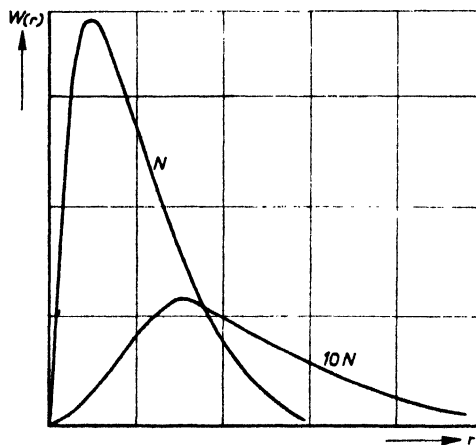
Fig. 5. The function $W(r)$ for $N = 100$ and $N = 1000$.

---

[1] E. GUTH and H. MARK, *Monatsh.*, 65 (1935) 93; RAYLEIGH, *Scientific Papers*, Vol. VI, p. 600.
[2] W. KUHN and F. GRÜN, *Kolloid-Z.*, 101 (1942) 248.
[3] Strictly speaking the integration should extend from 0 to NA, since NA is the largest possible value of r. However, since $W(r)$ decreases rapidly at high values of r, the upper limit may be replaced by ∞.

to the *average* value of r. In fact, it is easily derived from equation (7) that

$$r_{aver} = \sqrt{\frac{8}{3\pi}}A\sqrt{N} \lesssim A\sqrt{N} \tag{8}$$

This shows that the average length of the molecule for large values of N is very much smaller than the maximum length NA: this is the reason for the restriction imposed upon equation (4), where it was explicitly assumed that r was small compared with NA.

Similarly, the mean value of $1/r$ is proportional to $1/A\sqrt{N}$; that of r^3 is proportional to $A^3N\sqrt{N}$, and so on. Since the average value of r^3 is a measure for the volume occupied by the skein, we arrive at the important result that the average *density* in the skein is proportional to $A^{-3}N^{-\frac{1}{2}}$

If, for instance, the molecules have a slight tendency to associate, this association may also occur between different parts of a single molecule. Since the average density in the skein decreases with increasing N, this association will be *relatively less effective* in chains of greater length. Further, since for a given number of monomeric groups (i.e., for a given degree of polymerisation) $A^3\sqrt{N}$ will increase with increasing number ν of monomeric groups in a chain-element, the average density in the coil will be smaller in, say, cellulose derivatives, where ν is large, than in caoutchouc or paraffins, where ν is small. The importance of this conclusion will become clear in Section 4 and 6.

c. Final remarks

Let us now return to equation (6) giving the frequency distribution of the length r. Since Ψ_o is proportional to the number of ways in which a distance r may be arrived at, $k\ln\Psi_o$ may be said to represent the entropy s of a molecule with length r. (Compare the chapter on thermodynamics, p. 66; k is BOLTZMANN's constant). In other words

$$s = \text{const} - {}^3/_2\frac{k}{NA^2}\, r^2 \tag{9}$$

As it has been assumed throughout our treatment, that change of shape does not involve any change in potential energy, the free energy of the molecule becomes:

$$f = \text{const.} - {}^3/_2\frac{kT}{NA^2}\, r^2 \tag{10}$$

This interpretation implies, that a "force"

$$K = -\frac{\delta f}{\delta r} = \frac{3}{\,}\frac{kT}{NA^2}\, r \tag{11}$$

Fig. 6. "Force" K, driving one end of the molecule towards the other end.

is constantly driving one end of the molecule towards the other end. This force is counteracted by the BROWNean movement[1], thus giving rise to the equilibrium distribution (6). The situation is in complete analogy with the diffusion equilibrium of a large number of particles, moving in a field of force: the diffusion tends to bring about a uniform distribution; it is balanced by the forces acting on the particles (see chapter on thermodynamics, p. 66).

Finally, it should be emphasized that the theory can only be regarded as a first

[1] J. J. HERMANS, *Kolloid-Z.*, 103 (1943) 210.

approximation. It does not take into account that a chain-element cannot be present at points occupied by other ones; i.e., it neglects the finite size of the chain-elements. The correction required here would be similar in nature to VAN DER WAALS' *b*-correction in the theory of gases. No more did we take into account the possible forces acting between the chain-elements (VAN DER WAALS' *a*-correction). We have already incidentally touched upon the possibility that certain parts of the molecule would tend to associate with other parts. We do not believe that one need take into account a pronounced tendency towards association *all along the chain*, for the simple reason that if such tendency existed the molecules would not go into solution at all. (See chapter on thermodynamics p. 71). But if groups are present which show a quite specific interaction, neglect of this interaction may become very serious. It is obvious, for instance, that long-chain molecules (such as may be encountered in bio-colloids) carrying a large number of electric charges, will tend to assume such a form that the charges are lying at the periphery of the molecule. Similarly, if charges of opposite sign are present, these may easily give rise to internal associations[1]. Thus, the globular shape of many proteins in aqueous solution is due to specific interactions between their constituent parts. As was shown by LUNDGREN[2], such proteins may sometimes be made to unfold by the addition of certain substances. Although no theory regarding an interaction of the chain-elements is available at the present time[3], these effects should be a warning not to apply the results obtained to cases where their applicability may be doubtful.

§ 3. DIPOLE MOMENT OF LONG-CHAIN MOLECULES

One of the most convincing arguments in favour of the flexibility of long-chain molecules is furnished by their dipole moment. According to the well-known formula of DEBYE, the contribution of n dipoles of magnitude μ_o to the electric polarisation is

$$p = \frac{4}{3}\pi\frac{n\,\mu_o^2}{3\,kT}$$

Consequently, if a molecule carries n dipoles which do not influence each other's orientation in the external field, this molecule behaves like a dipole of magnitude

$$\mu = \mu_o \sqrt{n} \tag{12}$$

If applied to a macromolecule, this obviously means that its dipole moment will be proportional to $\sqrt{M}$, where M is the molecular weight. As was shown by BRIDGMAN[4], this is completely borne out by experiment. (Compare Fig. 7.) A specially simple case is that of long-chain molecules carrying two equal dipoles, one at each end. This case was examined by SMYTH and WALLS[5], who obtained complete agreement between theory and experiment, as is shown in the following table:

[1] Compare the chapter on colloids with electrolyte character.
[2] H. P. LUNDGREN, *J. Am. Chem. Soc.*, 63 (1941) 2854. See further: H. P. LUNDGREN, *Text. Res. J.*, 15 (1945) 335.
[3] An interesting attempt to extend the theory to zwitterions was made by W. KUHN, *Z. physik. Chem.*, A. 175 (1935) 1. See further the method outlined by H. EYRING, *Phys. Review*, 39 (1932) 746, and by H. A. KRAMERS, *Physica*, 11 (1944) 1.
[4] W. B. BRIDGMAN, *J. Am. Chem. Soc.*, 60 (1938) 530.
[5] C. P. SMYTH, W. S. WALLS, *J. Am. Chem. Soc.*, 53 (1931) 527, 2115.

TABLE 1
DIPOLE MOMENTS IN DEBYE UNITS

Substance	group moment μ_0	total moment (exper.) μ	$\mu_0 \sqrt{2}$
HO(CH$_2$)$_{10}$OH	1.65	2.4	2.3
Br(CH$_2$)$_{10}$Br	1.8	2.5	2.5
H$_5$C$_2$OOC(CH$_2$)$_8$ COOC$_2$H$_5$	1.8	2.5	2.5
H$_5$C$_2$OOC(CH$_2$)$_{10}$COOC$_2$H$_5$	1.8	2.5	2.5

A number of cellulose derivates in various solvents was examined by SAKURADA and LEE[1], who also studied polychloroprene, polyvinyl-acetate and polystyrene in benzene. In these cases the molecular weights were un-known, but it was shown by the following simple argument that the results obtained con-form to equation (12). If the concentration is c grams/cm^3, the solution contains

$$G = N_A \frac{c}{M} \qquad (13)$$

Fig. 7. Dipole moments of hydroxy-decanoic acids HO [(CH$_2$)$_9$COO)] $_{n-1}$(CH$_2$)$_9$ COOH.

molecules per ml, where M is the molecular weight and N_A is AVOGADRO's number. With a dipole μ per molecule, the polarisation per unit of volume becomes

$$p_0 = \frac{4}{3} \pi \frac{N_A c}{M} \frac{\mu^2}{3 \, kT} \qquad (14)$$

Now SAKURADA introduces the "dipole moment μ_g per monomeric group" ac-cording to the equation:

$$\mu^2_g = \frac{1}{P} \mu^2 = \frac{M_g}{M} \mu^2$$

If the dipole μ of the molecule is proportional to $\sqrt{M}$, the quantity μ_g must be independent of molecular weight. Now, from equation (14) it is seen that

$$\mu^2_g = M_g \frac{p_0}{c} \frac{3 \, kT}{\frac{4}{3} \pi N_A} \quad,$$

[1] I. SAKURADA and S. LEE, Z. physik. Chem., B 43 (1939) 245; 71 (1935) 94; 82 (1938) 67, 72, 194.

the value of which is entirely determined by experiment. The quantuty μ_g calculated in this way is independent of the degree of polymerisation and its order of magnitude is that of the monomeric group moment itself. Moreover, the value for triacetyl-starch is identical with that of triacetyl-cellulose [1].

A special type of dipole molecule is the zwitterion, where one of the charges is carried by one end of the molecule and the opposite charge by the other end. It can be shown [2] that the average value of the dipole moment in this case should again be proportional to $|\overline{M}|$, as is borne out by experiment [3]. It would be expected that this rule would no longer apply if M became very large. For, if r becomes large, the charges carried by the molecule are screened off to a certain extent by the gegenions present in the solution.

In this particular case of zwitterions the influence of the COULOMB potential on the statistical shape of the molecule was studied by KUHN [2]. The result obtained by him indicates, however, that the correction involved does not invalidate the general conclusion that μ should be practically proportional to $|\overline{M}$.

Although all experiments mentioned here give strong support to the assumption of randomly kinked structures, yet it is important to realise that they do not prove that the molecules are really *flexible*. A completely *rigid* structure of randomly oriented dipoles would lead to a total moment of exactly same magnitude. However, if the macromolecule were rigid, and had to be orientated as a whole, this orientation would be much slower than that of ordinary molecules, since the friction opposing the rotation of the macromolecule is much larger. From the fact that no anomalous dispersion has been found [4], we may conclude that all dipoles are free to turn practically independently.

4. SEDIMENTATION VELOCITY OF LONG-CHAIN COMPOUNDS

If a macromolecule of molecular weight M is subject to a centrifugal field of magnitude g, it experiences a force

$$\frac{gM}{N_A} (1 - V\varrho) \qquad (15)$$

where N_A is AVOGADRO's number, V is the specific volume of the solute and ϱ the density of the solvent [5]. This force will drag the molecule through the liquid with a velocity

$$u = \frac{g(1 - V\varrho)}{N_A} \frac{M}{w} \qquad (16)$$

where w is the frictional constant, i.e., the resistance experienced by the molecule when moving with unit velocity. Now, if we accept the picture developed in Section 2, our first impulse would be to assume that w is proportional to M, meaning thereby

[1] I. SAKURADA and S. LEE, *Kolloid-Z.*, 82, (1938) 67, 72.

[2] W. KUHN, *Z. physik. Chem.*, A. 175 (1935) 1.

[3] J. WYMAN and T. L. MC. MEEKIN, *J. Am. Chem. Soc.*, 55 (1933) 908, 915; G. DEVOTO, *Z. Elektrochem.*, 40 (1934) 490; W. KUHN and H. MARTIN Ber., 67 (1934) 1526.

[4] W. B. BRIDGMAN, *J. Am. Chem. Soc.* 60 (1938) 530.

[5] THE SVEDBERG and K. O. PEDERSEN, *Die Ultrazentrifuge*.

that a loosely built skein consisting of a number of particles would experience a resistance which was simply proportional to this number. If this were true, it would follow from equation (16), that the sedimentation velocity u would be independent of molecular weight. The experiments show, however, that u distinctly increases with increasing molecular weight [1, 2]. Several authors [3] have tried to explain this by assuming that the macromolecule behaves like a rigid ellipsoid or thin rod. In recent work on this subject [2, 4], however, a quantitative explanation could be given on the basis of a randomly twisted structure, taking into account the hydrodynamic interaction of its constituent parts. The following rough outline of this theory may suffice.

According to the well-known formula of STOKES, the resistance experienced by a sphere with diameter A, moving in a liquid of viscosity η_0 is

$$F = 3 \pi \eta_0 A v \tag{17}$$

where v is the velocity. It is further shown in STOKES' treatment, that the velocity of the liquid in the neighbourhood of the sphere decreases very slowly with increasing distance from the centre [5]. At a distance r from the centre the velocity is of the order of magnitude vA/r. Thus, if a second sphere is moving at a distance r from the first, its velocity will be increased by an amount of this order [6].

Now let us apply this result to the N statistical chain-elements composing the macromolecule. If they were all moving independently of each other, they would acquire the velocity

$$v = \frac{F}{3 \pi \eta_0 A}, \tag{18}$$

if F is the force acting on a chain-element. Any of the other chain-elements present produces an additional velocity of the order vA/r. The average contribution of these chain-elements to the velocity of the particular chain-element considered will be vA times the average value of $1/r$. In Section 2, p. 99 it was shown that this average value is proportional to $1/A\sqrt{N}$. The average contribution of a chain-element is therefore $\infty \ v/\sqrt{N}$. Since there are N of these chain-elements, [7] the total velocity becomes, instead of (18):

$$u \cong v + v\sqrt{N} = \frac{F}{3 \pi \eta_0 A} (1 + \sqrt{N}). \tag{19}$$

Obviously, if N is large, the hydrodynamic interaction will be very pronounced, and the velocity of the molecule as a whole becomes much larger than that of an isolated chain-element. The liquid in the immediate neighbourhood of the chain-elements is carried along almost entirely: in fact, with very large values of N, the

---

[1] THE SVEDBERG and K. O. PEDERSEN, *Die Ultrazentrifuge.*
[2] W. KUHN and H. KUHN, *Helv. Chim. Acta,* 26 (1943) 1394.
[3] R. SIGNER in SVEDBERG's book, p. 389, 392, 938; H. MOSIMANN, *Helv. Chim. Acta,* 26 (1943) 61.
[4] J. J. HERMANS, *Rec. trav. chim.,* 63 (1944) 219.
[5] This only applies to values of r which are not too large compared with A, since otherwise the inertia terms in the hydrodynamic equations begin to play a part. Compare C. W. OSEEN, *Hydrodynamik,* Leipzig 1927.
[6] J. M. BURGERS, *Proc. Akad. Wetenschappen Amsterdam,* 44 (1941) 1045, 1177; 45 (1942) 9, 126.
[7] Strictly speaking this number is $N - 1$, but the error is small if N is large.

velocity becomes [1] $F \vert \overline{N}/3 \pi \eta_o A$, which is nothing but STOKES' velocity for a sphere with diameter $A\sqrt{N}$ moving under the influence of a force NF. It may be mentioned in this connection that the hydrodynamic interaction described here is a well-known effect in meteorology, where it may explain the rapid motion of a fog or smoke [2].

It is worth noting that the correcting factor $1 + \vert \overline{N}$, for a given degree of polymerisation will be the larger, the larger N becomes, i.e., the smaller is the number ν of monomeric groups in a chain-element. This is due to the fact that, if ν is small, the density in the skein is high (compare Section 2 b, p. 99) and consequently the hydrodynamic interaction is more pronounced.

By way of illustration, table 2 gives the sedimentation constant of cellulose nitrate in acetone, as measured by MOSIMANN [3]. This sedimentation constant is the velocity in a centrifugal field of unit magnitude. The measurements extend over an exceptionally large region of molecular weights, viz., from $M = 6\,200$ to $M = 613\,000$. To calculate the value of N for a given value of M, one should of course, know the number of monomeric groups in a chain-element. As will be shown in the following section, this number can be derived from viscosity measurements. Doing this, we arrive at the third row in Table 2, which shows that σ is actually proportional

TABLE 2

SEDIMENTATION CONSTANT σ OF CELLULOSE NITRATE IN ACETONE

M = molecular weight, N = number of (statistical) chain-elements.

$10^{-3}M$	613	199	80.2	30	6.2
$10^{13}\sigma$	30	18	12.0	8.7	5.2
$10^{13}\sigma/(1 + \vert \overline{N})$	2.4	2.4	2.4	2.5	2.5

to $1 + \vert N$. It could further be shown [4] that the absolute value of σ also conforms to equation (19). A more detailed theory was developed by DEBYE[5] and independently, by BRINKMAN [6]. These authors considered more closely the flow of liquid through the sedimenting coil and derived a formula for the sedimentation velocity which shows that in the limit of very low molecular weight, M, this velocity is independent of M, whereas it becomes proportional to $\vert \overline{M}$ at higher values of M. For further details we refer to section 6c (p. 113) of this chapter.

§ 5. TIME OF RELAXATION

Some important properties of long-chain compounds result from reactions brought about by a deformation of the molecules. We know from Section 2 that the distribution ψ_o implied in equation (6) p. 98 may be regarded as a diffusion equilibrium between the "force" K on the one hand and the tendency towards diffusion on the

[1] W. KUHN and H. KUHN, Helv. Chim. Acta, 26 (1943) 1394.
[2] J. M. BURGERS, Proc. Akad. Wetensch. Amsterdam, 44 (1941) 1045, 1177; 45 (1942) 9, 126.
[3] H. MOSIMANN, Helv. Chim. Acta, 26 (1943) 61.
[4] J. J. HERMANS, Rec. trav. chim., 63 (1944) 219.
[5] P. DEBYE, XIth Intern. Congress of Pure and Appl. Chem., London, July 1947.
[6] H. C. BRINKMAN, Proc. Acad. Amsterdam, 50 (1947) 618.

other hand. Any action which results in disturbing this equilibrium distribution will bring forces into play.

First consider the simplest case, where the molecules are deformed and are then left to themselves [1]. The number of molecules with length r is now different from the expression ψ_o in equation (6). Consequently there will be a tendency to reestablish the equilibrium distribution ψ_o. This requires a certain time, the relaxation time Θ. In a solvent of viscosity η_o this relaxation time is

$$\Theta \sim \frac{\eta_o \, N^2 A^3}{kT}. \tag{20}$$

We shall not give a complete derivation of this formula since the nature of this derivation can be made clear by the following simple argument. In order to bring about a change in the length r, one end of the molecule must be moved with respect to the other, so causing a rearrangement of the chain-elements. The cause of this rearrangement is the force K, connecting the two ends of the molecule. From equation (11) we know that K is proportional to kT/NA^2. The velocity of the rearrangement as a result of this force is proportional to K/w, where w is the frictional constant, i.e., the resistance at unit velocity. Using STOKES' formula, each of the chain-elements will contribute an amount of the order of $3\pi\eta_o A$ to this frictional constant. Thus w will be proportional to NA, which means that the velocity K/w is proportional to $kT/\eta_o N^2 A^3$. Since the time of relaxation Θ will be inversely proportional to the velocity of rearrangement, we arrive at equation (20). In this reasoning we have neglected the hydrodynamic interaction of the chain-elements mentioned in the preceding section. (See below, Section 6.c, p. 113).

As $N = P/\nu$, where P is the degree of polymerisation and ν the number of monomeric groups in a chain-element, the relaxation time Θ will be proportional to P^2. Substituting $\eta_o \simeq 0.01$, which is the usual value for most solvents, and choosing $\nu \simeq 5$ as an average value for a number of compounds, we get

$$\Theta \simeq 5.10^{-12} P^2 \tag{21}$$

The result obtained will enable us to derive in a very simple way the contribution of the long-chain molecules to the viscosity of the solution. It is interesting to observe, however, that the relaxation time considered here may have a direct bearing on the phenomenon of depolymerisation by ultrasonics. We shall return to this subject in Section 9.g p. 122, since the depolymerising action of ultrasonic waves seems to be restricted to concentrations where the macromolecules are building a more or less coherent structure in the solution. For the present purpose the following remarks may suffice. From the fact that the macromolecules are broken down, SCHMID [2] infers that the structure is to be considered as a rigid one. The molecules obviously cannot follow the rapid oscillations in the sonic field. It would then be expected, however, that the effect would be restricted to frequencies surpassing the reciprocal relaxation time Θ. For, if Θ were much smaller than the period of an oscillation, the molecule would have no difficulty in following the sound waves. It is interesting to note in this connection, that polystyrene in toluene [3], if exposed to ultrasonic waves of frequency $2.8 \cdot 10^5$, is rapidly depolymerised to a degree of polymerisation of the order of 10^4, which is followed by a slower disintegration to about $3 \cdot 10^3$, where the process practically stops altogether. Substituting the value of 3.10^3 for P in equation (21), we find a relaxation time of about 5.10^{-6} seconds which is actually not far from the reciprocal frequency applied.

_____

[1] J. FRENKEL, Acta Physicochim. URSS, 9 (1938) 235; J. J. HERMANS, Kolloid-Z. 103 (1943) 218; 106 (1944) 24.

[2] G. SCHMID, Physikal., 41 (1940) 326; H. FREUNDLICH and D. W. GILLINGS, Trans. Faraday Soc., 205 (1938) 649.

[3] G. SCHMID, E. BEUTENMÜLLER, and A. RIEF, Kunststoff-Technik, 13 (1943) 65.

§ 6. VISCOSITY

a. Derivation of viscosity formula

The viscosity of long-chain compounds in dilute solution has been discussed in great detail on the basis of thin rods or ellipsoids[1]. It falls outside the scope of this chapter to enter into the subject. For particulars the reader is referred to BURGERS' article[2]. The theory shows that the contribution of a thin rod to the viscosity of the solution is roughly proportional to the third power of its length. Hence, if a macromolecule behaved as a rigid rod, the viscosity would be roughly proportional to P^3, where P is the degree of polymerisation. From STAUDINGER'S experimental rule, however, it can be concluded that the viscosity contribution is proportional to P^2, (see below), while the deviations from this rule point to an exponent which is somewhat smaller than 2, rather than larger[3].

HUGGINS[4] has given a derivation of STAUDINGER'S rule on the basis of randomly kinked structures. In this model, however, these structures are considered as *rigid*. In conformity with the treatment adopted in this chapter, we will here examine the hydrodynamic properties of the *flexible* chain[5].
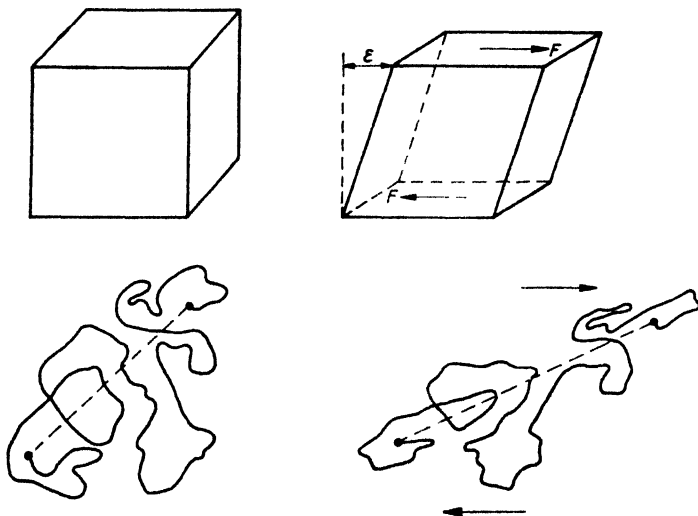
Fig. 8.   Deformation as a result of shear.

In the preceding section we examined only the laws governing the reestablishment of the equilibrium distribution. The nature of the deformation as a result of which this equilibrium distribution was upset is irrelevant to the process of recovery. We will now focuss our attention on a very special deformation of the molecules, viz.,

---

[1] Litterature can be found in P. H. HERMANS, J. J. HERMANS, and D. VERMAAS, *Kolloid-Z.*, 105 (1943) 199 or in W. PHILIPPOFF, *Viskosität der Kolloide*, Dresden and Leipzig 1922, p. 364 More recent contributions to the subject have been given by W. KUHN and H. KUHN, *Helv. Chim. Acta*, 28 (1945) 97.

[2] J. M. BURGERS, *Second Report on Viscosity and Plasticity*, Amsterdam, 1938.

[3] R. HOUWINK, *J. prakt. Chemie*, (2) 157 (1941) 15.

[4] M. L. HUGGINS, *J. Phys. Chem.*, 42 (1938), 911; 43 (1939), 439. Earlier attempts were made by H. MARK, *Z. Elektrochem.*, 40 (1934) 449; I. SAKURADA, *Z. physik. Chem.*, B 38 (1938) 407; W. KUHN, *Kolloid-Z.*, 68 (1934) 1; *Angew. Chemie*, 49 (1936) 858; See further W. HALLER, *Kolloid-Z.*, 56 (1931) 257.

[5] Compare also F. H. MÜLLER's tentative calculations in *Die Chemie, Beihefte* 47 (1943) 81

that brought about by a simple shear in the liquid. Fig. 8 gives a schematic representation of the situation encountered. The molecules, which are stretched by the shear imposed on the liquid, tend to reassume their original shape. This means that an extra resistance must be overcome, resulting in an increase in viscosity. Using a simile on a macroscopic scale, we may visualize the role of the molecules as that of elastic strings connecting the successive layers in the liquid. The resistance experienced by these strings when two layers are sliding past each other represents an extra link between the layers and thus contributes to the viscosity of the liquid. An expression for this viscosity may be derived from MAXWELL'S equation [1]

$$\eta = \theta \chi \qquad (22)$$

which relates the viscosity to the relaxation time θ and the shear modulus χ. Here χ represents the contribution of the chain-molecules to the force per unit area, assuming that the shear is performed within a time much shorter than θ. In other words, we consider a hypothetical deformation of such great rapidity that the molecules have no time to relax to any appreciable extent.

MAXWELL's relation is based on the following argument. In a solid body a finite deformation r of the type considered in Fig. 8 can only be maintained if a stress is applied (symbolised by the arrows F). The force per unit area needed to maintain unit shear is called shear modulus χ. In other words

$$F = \chi r \qquad (23)$$

If a similar deformation is imposed on a liquid, the stress will soon disappear, since the molecules in the liquid are free to yield to the forces acting on them: the liquid flows. MAXWELL assumes that the stress will die down according to a simple relaxation process:

$$\frac{-dF}{dt} = \frac{F}{\theta} \qquad (24)$$

where θ represents the time of relaxation. Now assume that the shear r is increasing continuously (the liquid is flowing). Then the stress F increases according to equation (23) and simultaneously decreases according to equation (24). In other words, it is assumed that the two processes may be simply superimposed:

$$\frac{dF}{dt} = \chi \frac{dr}{dt} - \frac{F}{\theta} \qquad (25)$$

Applying this equation to continuous flow, where the stress is constant, we find that the velocity dr/dt as a result of a constant stress F is given by

$$\frac{dr}{dt} = \frac{F}{\theta \chi}$$

On account of the definition of viscosity, this means that equation (22) applies.

The contribution of the macromolecules to the modulus of shear can be derived from considerations involving the entropy change at deformation[2]. We shall not reproduce this derivation here but mention only the result:

$$\chi = kTG \qquad (26)$$

where G is the number of molecules per unit of volume. This expression can be made plausible by the following simple argument. In the first place, the contribution

[1] W. KUHN, *Naturwiss.*, 26 (1938) 661; *Z. physik. Chem.*, B 42 (1939) 1.
[2] J. J. HERMANS, *Kolloid-Z.*, 106 (1944) 22. Compare similar calculations by W. KUHN, *Kolloid-Z.*, 68 (1934) 2; 76 (1936) 258.

of a molecule to the shear modulus will be proportional to the force $K = 3kTr/NA^2$, connecting the two ends (compare equation (11) on p. 99). Further, if we divide the molecules into groups according to their length r, it is clear that the contribution of such a group to the shear modulus will be proportional to r, since the molecules with large values of r on the average are connecting two layers which are lying at a greater distance from each other. Thus the average contribution is proportional to

$$\frac{kT}{NA^2}\ \overline{r^2}$$

where $\overline{r^2}$ represents the mean value of r^2. As was shown, however, in section 2b., this mean value is proportional to NA^2, and thus, multiplying by the number G of molecules per unit of volume, we get equation (26).

In view of equations (20) and (22) it is obvious that the contribution of the solute to the viscosity is proportional to

$$\eta' \infty\ \eta_0\, N^2 A^3 G \tag{27}$$

This is the extra viscosity produced by the chain-molecules. Expressing it in terms of *specific viscosity increase* $\triangle\eta_{sp} = \eta'/\eta_0$, we get

$$\triangle\eta_{sp} \infty\ N^2 A^3 G \tag{28}$$

There exists a slightly different approach to the problem, consisting of calculating the distribution function in the streaming liquid[2]. As a result of the velocity gradient one end of the molecule is continuously carried away from the other end: it is clear that the velocity in the upper part of the skein (Fig. 8) relative to the molecule as a whole is opposite to that in the lower part. In this way a new distribution ψ is generated, which differs from the equilibrium distribution ψ_0 in equation (6). This method leads to the same value of $\triangle\eta_{sp}$.

The formula (28) is essentially the same as that obtained by HUGGINS[1] on the basis of rigid instead of flexible coiled structures. This is due to the fact that at low values of shear-rate, the deformation of the molecule is small and its viscosity contribution consequently negligibly different from that of the undeformed molecule. At higher rates of shear, however, HUGGINS introduced a correction for flexibility, whereas, according to our treatment, the result (28) is independent of the velocity gradient (see also p. 113).

b. Interpretation of experimental data

The equation (28) is in conformity with STAUDINGER's viscosity rule[3], according to which

$$[\eta] = \left[\frac{\triangle\eta_{sp}}{c}\right]_{c=0} = K_\eta\, P, \tag{29}$$

where c is the concentration by weight and K_η is a constant. In fact, equation (28)

[1] M. L. HUGGINS, *J. Phys. Chem.*, 42 (1938) 911; 43 (1939) 439; *J. appl. Phys.*, 10 (1939) 700.
[2] W. KUHN and H. KUHN, *Helv. Chim. Acta*, 26 (1943) 1394; J. J. HERMANS, *Physica*, 10 (1943) 777.
[3] H. STAUDINGER, *Organ. Koll. Chemie*, Braunschweig 1940, p. 54, 141, 163.

states that the contribution of a molecule with degree of polymerisation P is proportional to P^2. Since the concentration c, for a given number G is proportional to P, (28) is equivalent to (29). Following the terminology introduced by KUHN[1], K_η will be called the "viscosity number". From equation (28) it is apparent that this viscosity number contains the quantity ν. Since $N = P/\nu$, and since A is roughly proportional to ν, K_η will be the larger the larger is ν. Conversely, the value of ν may be estimated from viscosity experiments. This has been done for a few substances in Table 3, using STAUDINGER'S[2] values of K_η. The number of monomeric groups

TABLE 3

THE NUMBER ν OF MONOMERIC GROUPS IN A CHAIN-ELEMENT, AS DERIVED FROM VISCOSITY

Substance	Monom. group	solvent	K_η	ν
paraffin	$-CH_2-$	benzene	0.093	4—5
caoutchouc	$-CH=CCH_3-CH_2-CH_2-$	toluene	$\pm$ 0.17	4—6
polystyrene	$-CHC_6H_5CH_2-$	benzene	0.18	7—9
cellulose triacetate	$C_6H_7O_2(OCOCH_3)_3$	chloroform	0.53	∞12
,, ,,	,, ,,	m-cresol	0.63	∞16
cellit	$\pm$ $C_6H_6O_3(OCOCH_3)_2$	acetone	0.90	∞18
methylcellulose	$\pm$ $C_6H_8O_3(OCH_3)_2$	water	1.1	∞22

in a chain-element, obtained in this way is, on the whole, quite reasonable. It is obvious that we must be content with a rough estimate, because the linear dimensions A of the chain-element play a part in the result (compare eq. (28). Table 4, for instance, shows the influence of a substituent on the hydrodynamic properties of a chain. This influence may be attributed to a change in the flexibility of the chain

TABLE 4

CELLULOSE NITRATE OF VARYING NITROGEN CONTENT IN ACETONE[3]

percentage nitrogen.	13.79	13.60	12.75	12.11	10.70
NO_2 groups per C_6	2.9	2.8	2.5	2.2	1.8
value of K_η (c in g/ml).	0.177	0.174	0.130	0.116	0.097
corresponding ν-value	28	27	19	16.5	13

(compare the value of ν in Table 4), but it is obvious that volume effects may also enter into the matter, especially if the NO_2-groups have a tendency towards solvation.

In this connection the influence of the solvent on the specific viscosity is also a matter of great interest. An example is given in Table 4. Here again any interpretation is as yet highly speculative[4]. STAUDINGER believes[5] that $\triangle \eta_{sp}$ is large in "good" solvents, where the solvation is presumably more pronounced. SAKURADA[6],

[1] W. KUHN and H. KUHN, Helv. Chim. Acta, 26 (1943) 1394.
[2] See note 3 page 108.
[3] H. A. WANNOW, Kolloid-Z., 102 (1943) 29.
[4] The subject has been treated in some detail by H. A. STUART, Naturwiss., 31 (1943) 123, and by A. LÜTTRINGHAUS, Naturwiss., 30 (1942) 40.
[5] H. STAUDINGER, Berichte, 63 (1930) 2317.
[6] I. SAKURADA and M. SHOJINO, Kolloid-Z., 68 (1934) 300.

TABLE 5

CAOUTCHOUC IN VARIOUS SOLVENTS [1] (C IN GRAMS PER LITRE)

solvent	$\triangle \eta_{sp}/c$
petr. ether	5.05
$CHCl_3$	5.85
tetralin	5.92
C_6H_5Cl	6.21
CS_2	6.33
C_6H_6	6.85
CCl_4	7.94

however, could not confirm STAUDINGER'S results in this respect. The problem was further approached by FRITH[2], who derived a relation between specific viscosity increase and molecular interaction. The more energy is needed to bring the polymer molecules into solution, the greater is the tendency to form polymer-polymer contacts in the solution. We may expect, therefore, that the degree of coiling is to some extent dependent on solvent-solute interaction. This will evidently affect the viscosity. HUGGINS [3] has used this effect to explain the influence of solvent type on the value of $[\eta]$, while FRITH's theory relates to the slope of the η vers. c curve rather than to the limiting value $[\eta]$.

From equation (28) it is further apparent, that the specific viscosity is independent of the temperature in so far as ν and A may be considered as constants. Experiments of SAKURADA and collaborators[4], MEYER and VAN DER WYK[5] and DANES [6] have shown that $\triangle \eta_{sp}/c$ decreases if the temperature is increased. This decline is of the order of 1% per degree. It shows that the flexibility of the molecule is increased, as would be expected. It is interesting to note also that the dipole moment of macromolecules is usually sensitive to temperature changes [7]. The relative change in this dipole moment is of the same order of magnitude as that in $[\eta]$.

In other cases, however, the phenomena cannot be explained in this simple manner. In fact, the degree of coiling depends, among others, on the interaction of the chain-elements with the solvent molecules. In other words, the heat of mixing is involved. The influence of the temperature on $\triangle \eta_{sp}$ will therefore depend on this heat of mixing.

c. Validity of STAUDINGER'S rule

As regards the validity of STAUDINGER's viscosity rule, opinions are still at variance [8]. It is beyond all doubt that a number of cellulose derivatives conform closely

[1] I. SAKURADA et al. J. Soc. chem. Ind. Japan, 37 (1934) 468, 470, 486.
[2] E. M. FRITH, Trans. Faraday Soc., 41 (1945) 17, 90. Experimental work by ALFREY, BARTOVICS, and MARK, J. Am. Chem. Soc., 64 (1942) 1557.
[3] M. L. HUGGINS, J. appl. Phys., 14 (1943) 246; see also A. G. JANSSEN and B. P. CALDWELL, Polymer Bull., 1 (1945) 120.
[4] I. SAKURADA, J. Soc. chem. Ind. Japan, 37 (1934) 468; I. SAKURADA and K. TANAKA, ibid., p. 470; M. TANIYUCHI and I. SAKURADA, ibid., p. 486.
[5] K. H. MEYER and A. VAN DER WIJK, Kolloid-Z., 76 (1936) 287.
[6] V. Z. DANES, Kolloid-Z., 68 (1934) 110.
[7] I. SAKURADA and S. LEE, Kolloid-Z., 82 (1938) 67, 72.
[8] K. H. MEYER and A. VAN DER WIJK, Helv. Chim. Acta, 18 (1935) 1067; 19 (1936) 218; Z. Elektrochem., 40 (1934) 446; H. STAUDINGER, Helv. Chim. Acta, 19 (1936) 204; Wo. OSTWALD, Kolloid-Z., 106 (1944) 1.

to this rule[1]. Other substances however, clearly deviate (compare Fig. 9a). This is best shown by plotting K_η against degree of polymerisation (Fig. 9b).

Sometimes the descrepancies are reconciled by assuming a relation of the form

$$[\eta] = KP + K'$$

where K' is an additional constant[2] (usually negative). The theoretical meaning of this constant must perhaps be sought partly in changes in the liquid structure of the solvent brought about by the presence of the solute molecules.

In many cases, however, there is clearly no strictly linear relation between $[\eta]$ and P. MARK and others[3] have discussed a formula of the type

$$\triangle \eta_{sp}/c = KP^n, \tag{30}$$
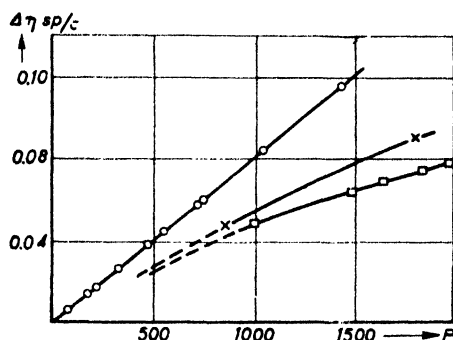
Fig. 9a

$\triangle \eta_{sp}/c$ plotted against P (c in g/liter)
○   cellulose nitrate in acetone[1].
×   polymethacrylicnitrile in acetone[4]
□   polyvinylacetate in acetone[5]
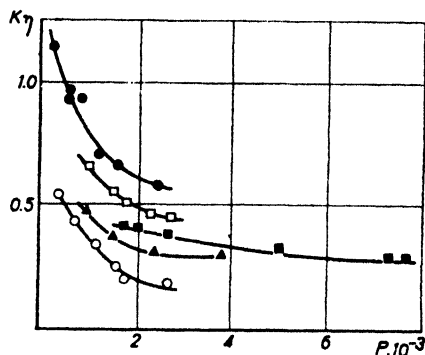
Fig. 9b

$K_\eta = \triangle \eta_{sp}/cP$ plotted against P
○   polymethyl methacrylates in acetone.
●   polymethyl methacrylates in $CHCl_3$.
▲   polymethyl acrylate in acetone.
■   polyvinalacetate in acetone.
□   polyvinylalcohol in water.

where n may assume values between 0.5 and 0.9. We shall have to consider whether these deviations can be accounted for. Now it is clear, that no theory can claim general applicability. Our calculations did involve a number of approximations which might have different influences upon different substances. In particular, we know already that deviations are likely to occur if certain atoms in the macromolecule show a strong specific interaction (compare Section 2c of this chapter). In

[1] G. V. SCHULZ, J. makromol. Chemie (3) 1 (1944) 149.
[2] W. O. BAKER, C. S. FULLER, and J. H. HEISS, J. Am. Chem. Soc., 63 (1941); 2142; G. GEE, Trans. Faraday Soc., 40 (1944), 264.
[3] M. L. HUGGINS, J. appl. Physics, 14 (1943) 246; J. W. TAMBLYN, D. R. MOREY, and R. H. WAGNER, Ind. Eng. Chem., 37 (1945) 573; M. FOURNIER and X. THIESSE, Compt. rend., 222 (1946) 1437; W. KUHN, Kolloid-Z., 68 (1934), 1; Angew. Chemie, 49 (1936) 858; I. SAKURADA, Z. physik. Chem., B 38 (1938), 407; A. MATTHES, J. Prakt. Chem. (2) 162 (1943) 273; F. H. MÜLLER, Beih. Die Chemie, 47 (1943), 81; R. HOUWINK, J. prakt. Chem. (2) 157 (1941) 15; and others.
[4] W. KERN and H. FERNOW, J. prakt. Chem., (2) 160 (1942) 307.
[5] H. STAUDINGER and H. WARTH, J. prakt. Chem. (2) 155 (1940) 278

this connection HUGGINS (*loc.cit.*) studied the relation between polymer-solvent interaction and the value of the exponent n in equation (30).

Moreover, we have neglected the hydrodynamic interaction of the chain-elements and we know from the discussion on sedimentation velocity that this hydrodynamic interaction may be a matter of considerable importance. In view of the results obtained in Section 4 we may even ask whether any weight may be attached at all to a method which does not take this interaction into account. The answer is that the hydrodynamic interaction in a velocity gradient is much smaller than if the molecule is moving as a whole, for the simple reason that the velocity in the upper part of the skein is opposed to that in the lower half. The forces arising from the flow of liquid along the chain-elements in the upper part are balanced to a certain extent by those generated in the lower part. This explains why there exist a number of substances which conform closely to STAUDINGER's rule while at the same time their sedimentation velocities distinctly increase with increasing molecular weight[1].

There can be no doubt, however, that the hydrodynamic interaction will also play an important part in a velocity gradient if the molecular weight is sufficiently high. This will apply the sooner the lower the number v of monomeric groups in a chain-element. For, if v is small, the density in the skein and therefore the hydrodynamic interaction of the chain-elements is large. In the limit of a completely matted coil we may expect that the molecule with the enclosed solvent will behave as a solid sphere. This would lead to a formula of the type $[\eta] = KP^{\frac{1}{2}}$. This model was worked out in full detail by SADRON[2]. According to BADGLEY[3], $[\eta]$ is proportional to P in the range of comparatively low molecular weight and becomes proportional to P^n ($n < 1$) for higher P-values. For tightly curled molecules n approaches the value $\frac{1}{2}$. This is exactly what should be expected on theoretical grounds. BADGLEY himself proposes a formula of the type

$$[\eta] = KP - K'P^2$$

We thus arrive at the important result that STAUDINGER's viscosity rule must be considered as a limiting law for very loose structures. The closer the chain-elements are packed in the skein, the sooner will deviations occur. This explains why the cellulose derivatives having a large v-value, are in particularly good agreement with STAUDINGER's rule. Here the hydrodynamic interaction is comparatively small. It is beyond our aim, however, to make a deep-searching study of the subject: too many factors would have to be considered. STAUDINGER[4] believes that in a number of cases his viscosity rule is invalidated as a result of ramification. Further, with many of the substances examined the determination of the molecular weight would have to be reconsidered very carefully, while moreover the macromolecules have not always been fractionated to a sufficient degree. It would lead us much too far to enter into any details, and we will therefore close the subject by suggesting that STAUDINGER's rule, although certainly not of general applicability, may serve, in principle, as a key to the interpretation of viscosity measurements.

    This statement is at variance with the opinion of some other authors. SADRON[2] in particular makes the objection that it is not permissible to assume that

[1] J. J. HERMANS, *Rec. trav. chim.*, 63 (1944) 219.
[2] CH. SADRON, *J. chim. phys.* 44 (1947); see also p. 103 of this chapter.
[3] W. J. BADGLEY, *Polymer Bull.*, 1 (1945) 17.
[4] H. STAUDINGER, *Organ. Kolloidchemie*, Braunschweig 1940.

all chain-elements are fully exposed to the hydronamic shearing forces. Rather are most of the chain-elements shielded by neighbouring chain-elements in the molecule. In other words, this author emphasizes the hydronamic interaction of the chain elements discussed by us on p. 103 and 112. To account for this effect, he introduced the so-called "equivalent" particle, whose surface may be indentified more or less with the envelope of the random coil. If the molecular weight is high this "equivalent" particle is a sphere, and the viscosity formula becomes identical with (30), with $n = \frac{1}{2}$ (completely matted coil). Whereas our treatment assumes a more or less independent action of the shear on the chain-elements present, and considers the hydrodynamic interaction as a complicating secondary phenomenon, SADRON's treatment represents the other extreme.

An answer to this problem, which accounts for the principal features of the phenomena, was given recently by DEBYE[1] and by BRINKMAN[2]. These authors considered the flow of liquid through a random coil in greater detail. The resistance experienced by the liquid as a result of the chain-elements present is accounted for by an additional term in the hydrodynamic differential equation. Solving the differential equations thus obtained, one finds the flow of liquid through the "porous structure" constituting the chain-molecule and from this the viscosity contribution of the molecule. The result obtained shows that STAUDINGER's rule should apply in the limit of very low molecular weight (always assuming that the coil has its random spherical shape). In this limit the coil is completely free-drained; the hydrodynamic interaction of the chain-elements is negligible. With increasing number of chain-elements $\triangle \eta_{sp}$ increases less rapidly than proportional to the molecular weight, M, and becomes proportional to $M^{\frac{1}{2}}$ for very high values of M. BRINKMAN shows that the relation between $\triangle \eta_{sp}$ and M for a considerable range of M-values may be approximated by MARKS's formula $\triangle \eta_{sp} = KM^n$, with $n = 0,71$. DEBYE considers in some more detail the change of the exponent n with molecular weight.

d. Non-Newtonian viscosity in dilute solution

Finally, a few words may be added with regard to the viscosity at high rates of shear. From the hydrodynamic theory of rod-shaped particles it can be inferred that the viscosity contribution of the rods is decreased as the rate of shear increases[3]. This is due to the fact that the particles are more and more oriented in the direction of flow. With flexible long-chain molecules this effect does not exist. It is true that here also the orientation is more pronounced at high rates of shear. This would result in a lower viscosity. At the same time, however, the elastic forces generated by the deformation of the molecules are increased. As the result of these two effects the vis-cosity is independent of the rate of shear.

This seems to be confirmed by experiment[4]. We mention in particular LYONS' measurements in cuprammonium solutions of cellulose[5]. According to this author the relative viscosity a as function of concentration c is best epresented by the formula

[1] P. DEBYE, *XIth Internat. Congress of Pure and Appl. Chem., London*, July 1947
[2] M.C. BRINKMAN, *Proc. Acad. Amsterdam* 50 (1947) 618; *Physica* 13 (1947) 447.
[3] A. PETERLIN, *Z. Physik*, 111 (1938) 232; J. R. ROBINSON, *Proc. Roy. Soc. London*, A 169 (1938) 156.
[4] F. H. MÜLLER, *Beih. Die Chemie*, 47 (1943) 81; W. O. BAKER, C. S. FULLER and J. H. HEISS, *J. Am. Chem. Soc.*, 63 (1941) 3324.
[5] W. J. LYONS, *J. Chem. Phys.*, 13 (1945) 43.

$$\frac{\eta}{\eta_o} = (1 + \frac{c}{\lambda})^8 + (K_i - \frac{8}{\lambda})c$$

where λ depends on the velocity gradient, but K_i does not, showing that lim $(\triangle \eta_{sp}/c)$ is independent of the shear.

It is to be noted, however, that the theory is likely to fail at very high rates of shear [1] and it would therefore be of great interest to re-examine this point in more detail.

In this connection we must mention the recent work of KUHN and KUHN [2]. We have seen on p. 94 that the freedom of rotation about single bonds in a chain-molecule is restricted by steric hindrance, and that potential barriers of considerable height must be overcome to bring about changes in shape. KUHN and KUHN have analysed this, focussing their attention on changes in the distance r between the two endpoints of the molecule. Changes in r can be brought about by rotation about single bonds from one equilibrium position into the neighbouring one. Since these positions are separated by a potential barrier, these "jumps" from one equilibrium position into the next are comparatively infrequent. For this reason changes in the distance r tend to be comparatively slow. In other words, the existence of the potential barriers means that there exists a resistance against rapid changes in shape. This effect can be accounted for by the introduction of an "inner viscosity" or "viscosity of form" of the chain-molecule. For a given type of polymer molecule calculations show this inner viscosity to be inversely proportional to the molecular weight.

It is assumed that the resistance due to inner viscosity can be simply superimposed onto that resulting from the viscosity of the medium surrounding the chain-molecule. Then, clearly, a molecule with high inner viscosity will behave as if it were almost rigid, whereas a molecule with small inner viscosity conforms to the laws derived in the preceding sections of this chapter. In particular, a molecule with small inner viscosity will give rise to a viscosity contribution $[\eta]$ which is independent of the rate of shear. A molecule with large inner viscosity, however, will show a behaviour which in certain respects is similar to that of a rigid rod. According to KUHN and KUHN [3], $[\eta]$ will decrease with increasing rate of shear, and the authors believe this to be confirmed by experiment. Conversely, the inner viscosity of the chain-molecules may be calculated from viscosity data.

Admittedly the height of the potential barriers calculated from this experimental "inner viscosity" tallies well with known data. Yet there are a number of objections to KUHN's conclusions. To begin with, the experimental difficulties in the measurement of $[\eta]$ as a function of shear are great, since it involves extrapolation to infinite dilution. We have already seen that the recent results obtained by LYONS for cellulose in cuprammonium did not reveal any change in $[\eta]$ with shear. Further it is obvious that the resistance to changes in shape experienced by the chain molecule as a result of inner viscosity and that resulting from the viscosity of the solvent are not simply additive. For this reason, however interesting KUHN's treatment, we are not yet in a position to consider his results as conclusive. Further studies in this interesting field will be of great value for the understanding of the physical properties of macromolecules.

[1] H. A. KRAMERS, Physica, 11 (1944) 1; J. J. HERMANS, Rec. trav. chim., 63 (1944) 205.
[2] W. KUHN and H. KUHN, Helv. Chim. Acta, 28 (1945) 1533; 29 (1946) 71, 609, 830.
[3] W. KUHN and H. KUHN, Helv. Chim. Acta, 28 (1945) 97.

§ 7. BIREFRINGENCE OF FLOW

a. Outline of the theory

The orientation of the molecules brought about by the velocity gradient is revealed by a double refraction. A detailed discussion of this double refraction on the basis of rigid ellipsoidal particles can be found in the work of PETERLIN and STUART [1]. In recent work [2] the subject was treated from the point of view of randomly coiled flexible structures. The principal features of this treatment may be outlined as follows.

From Section 2 we know that each molecule shows a double refraction, the magnitude of which is a function of the distance r between the two ends. If the solution is at rest, the distribution of the vectors r over all orientations in space is uniform; the liquid as a whole is optically isotropic. In a velocity gradient, however, the molecules besides being stretched are oriented in the direction of flow. As has already been mentioned in the preceding section, the statistical distribution of the orientation in the streaming liquid can be calculated. Using the formula (4) on page 97 for the anisotropy of a molecule, and averaging over all molecules in the streaming liquid, we arrive at an expression for the anisotropy of the solution. This procedure assumes that the molecules in the velocity gradient on the average retain their symmetry of rotation round the vector r. We shall not go into mathematical details, but shall mention only the result obtained. For the rest we refer to Volume I.

For definiteness let us assume that the experiment is carried out in a cylindrical COUETTE apparatus. This may be represented by Fig. 10 a, where the outer cylinder is assumed to rotate with respect to the inner one. If the radii of the two cylinders are sufficiently large compared with the distance between them, the streaming is practically identical with the simple laminar flow considered in Fig. 8. A beam of linearly polarised light travels through the solution in the direction of the z-axis, i.e., perpendicular to the plane of drawing in Fig. 10. It can be shown that there are three principal axes of polarisability in the streaming liquid. One of these axes, Z, coincides with the z-axis. The other two, X and Y, are shown in Fig. 10 b. These are the axes of largest and smallest refractive index. The angle Φ between x and Y

Fig. 10 a. Cylindrical COUETTE apparatus.   Fig. 10 b. Ellipsoid of polarisability in the streaming liquid.

---

[1] A. PETERLIN and H. A. STUART, *Hand- und Jahrb. Chem. Physik VIII*, Abschnitt I B, Leipzig 1943. See also W. KUHN, *Z. phys. Chem.*, A. 161 (1932) 1 427; *Kolloid-Z.*, 62 (1933) 269; W. HALLER, *Koll.-Z.*, 61 (1932) 26; P. BOEDER, *Z. Physik*, 75 (1932) 258.

[2] W. KUHN and H. KUHN, *Helv. Chim. Acta*, 26 (1943) 1394; J. J. HERMANS, *Physica*, 10 (1943) 777; H. A. KRAMERS, *Physica*, 11 (1944) 1. Review-article in W. KUHN, *Experientia*, 1 (1945) 6

is called extinction angle. The difference between the refractive indices n_Y and n_X for oscillations parallel to the Y and the X-axis respectively is called double refraction:

$$\triangle n = n_Y - n_X$$

The theory[1] leads to expressions for both $\triangle n$ and Φ, both of which can be determined experimentally.

b. Double refraction

The formula for $\triangle n$ is similar in nature to that for the viscosity contribution:

$$\triangle n = k_1 \, \eta_0 \, D \, \frac{a_1 - a_2}{kT} \, N^2 A^3 G \tag{31}$$

Here D represents the magnitude of the velocity gradient, while a_1-a_2 is the anistropy of the chain-element (compare Section 2). The factor k_1 is a simple function of the refractive index of the solvent and is irrelevant to the present argument. Formula (31) shows that the contribution of a molecule to the double refraction is proportional to P^2 if P is the degree of polymerisation. Thus $\triangle n/c$ is proportional to P, where c represents the concentration by weight. From analogy with the "viscosity number" mentioned in the preceding section, KUHN[2] introduced the "optical number" K_n, which is defined by the equation:

$$\frac{\triangle n}{\eta_0 D c} = K_n P \tag{32}$$

This rule was found experimentally by SIGNER[3]. It is completely analogous to STAUDINGER's viscosity rule and is subject to the same restrictions (compare the preceding section). The two equations (31) and (28) may be combined to give

$$\frac{\triangle n}{\eta . D} = k_1 \, \frac{a_1 - a_2}{kT} \, \triangle \eta_{sp} \tag{33}$$

From this equation the anisotropy a_1-a_2 of the chain-element may be calculated if $\triangle n$ and $\triangle \eta_{sp}$ have been measured. This leads to values which are of the same order of magnitude as the anisotropy of a monomeric group[4]. No accurate check is possible since the influence of the solvent on the magnitude of a_1-a_2 is unknown (compare Section 2a p. 96).

c. Extinction angle

The theoretical formula for the extinction angle Φ is of the form:

$$\operatorname{tg}2\Phi \sim \frac{kT}{\eta_0 D N^2 A^3} \tag{34}$$

It was written by KUHN and KUHN[5] in terms of the degree of polymerisation:

[1] W. KUHN and H. KUHN, Helv. Chim. Acta, 26 (1943) 1394; J. J. HERMANS, Physica, 10 (1943) 777; H. A. KRAMERS, Physica, 11 (1944) 1.
[2] W. KUHN and H. KUHN, Helv. Chim. Acta, 26 (1943) 1394.
[3] R. SIGNER, Z. phys. Chem., 150 (1930) 257; R. SIGNER and H. GROSS, Z. phys. Chem., A. 165 (1933) 161.
[4] J. J. HERMANS, Rec. trav. chim., 63 (1944) 25.
[5] W. KUHN and H. KUHN, Helv. Chim. Acta, 26 (1943) 1394.

$$\frac{1}{\eta_o Dtg\ 2\Phi} = 2K_\phi\ P^2 \tag{35}$$

where K_ϕ is called "orientation number". The proportionality of tg 2Φ with D and with $1/P^2$ is, on the whole, borne out by experiment[1,2]. It is obvious from equation (34) that there again the number $\nu = P/N$ may be calculated from experimental results. The value of ν so obtained tallies well with that calulated from viscosity determinations[2]. It is to be noted, however, that tg 2Φ is very sensitive to deviations from the average molecular weight (see Chapter on molecular weight determinations, p. 143). The substances investigated so far have not been fractionated to a sufficient degree to allow of a decisive check. As, moreover, experiments on the birefringence of flow in solutions of long-chain molecules are still scarce, and have usually been carried out at concentrations which are too high for the interaction of the solute molecules to be neglected, further work on the subject is desirable.

§ 8. KRAMERS' METHOD. RAMIFIED MOLECULES AND RINGS

It has been assumed so far that the behaviour of the molecules could be described with sufficient accuracy by the distance between the two ends. The fact that these ends are connected by a series of chain-elements was expressed by the introduction of the fictive force K [equation (11)]. The further development of the theory required some additional assumptions. We may mention, for instance, that the friction experienced in the motion of the two ends towards each other was assumed to be proportional to the total number N of statistical chain-elements, since on the average all these chain-elements will be involved. Further, a quantitative treatment of the birefringence of flow was only made possible by assuming that the molecules on the average retain their symmetry of rotation in the streaming liquid.

It was shown by KRAMERS[3] that the results obtained are confirmed by a more accurate theory. Instead of the laminar flow considered in Section 6, KRAMERS assumes a velocity distribution which is derived from a velocity potential. This means that the components $u,v,w,$ of the velocity in the liquid can be written as the derivatives of a function $U\ (x,y,z)$.

$$u = -\frac{\partial U}{\partial x} \qquad v = -\frac{\partial U}{\partial y} \qquad w = -\frac{\partial U}{\partial z} \tag{36}$$

For simplicity let us consider the equation of motion in one dimension (say x) only. The force X acting on a chain-element is the sum of the irregular force X', which results from the BROWNIAN movement in the liquid, and the frictional force $-\zeta\ (u_1 - u)$:

$$X = -\zeta\ (u_1 - u) + X' \tag{37}$$

Here ζ is a frictional constant, while u_1 is the velocity of the chain-element. In other words, u_1-u is the velocity of the chain-element with respect to its immediate surroundings. In view of equation (36) we may write

[1] W. KUHN and H. KUHN, *Helv. Chim. Acta*, 26 (1943) 1394.
[2] J. J. HERMANS, *Rec. trav. chim.*, 63 (1944) 25.
[3] H. A. KRAMERS, *Physica*, 11 (1944) 1.

$$X = -\frac{\partial U}{\partial x} - \zeta u_1 + X' \tag{38}$$

The equations of motion thus obtained are those which would apply if the liquid were at rest and the chain-elements subject to a field of force with potential energy. U. It is therefore possible to attack the problem with the general methods of statistical mechanics[1].

The potential U which is best adapted to the problem considered has the form

$$U = -\tfrac{1}{2} qxy \tag{39}$$

Thus $u = \tfrac{1}{2} qy, \quad v = \tfrac{1}{2} qx, \quad w = 0.$

This streaming represents a superimposition of the simple laminar flow

$$u = qy, \quad v = 0, \quad w = 0;$$

and the uniform rotation

$$u = -\tfrac{1}{2} qy, \quad v = \tfrac{1}{2} qx, \quad w = 0.$$

It can be shown that the influence of this rotation is confined to negligible terms which are quadratic in the rate of shear q.

It would lead us too far, however, to enter into the matter any further. For our purpose it suffices to know that the contribution of a molecule consisting of N chain-elements to both the viscosity and the birefringence is shown to be proportional to N^2, which is in complete agreement with the result obtained in this chapter. It may further be mentioned that KRAMERS also succeeded in computing the viscosity of molecules which possess branching points or rings. If, for instance, the molecule consists of s equal branches (each consisting of N/s chain-elements), which start from a common point, the viscosity is γ times that of the unbranched molecule consisting of N chain-elements. Here

$$\gamma = \frac{3}{s} - \frac{2}{s^2} \tag{40}$$

which equals unity if $s = 1$ or 2 as it should be, and which is always below unity if $s > 2$.

If the two ends of a molecule are connected to form a ring, the viscosity contribution is reduced by a factor 2.

§ 9.   CONCENTRATED SOLUTIONS

Throughout the preceding treatment it was explicitly assumed that the concentration of the solution was sufficiently small to disregard all interaction between the solute molecules. As soon as this condition is violated, conditions often become so complicated that a quantitative treatment is illusory at the present time.

a.   Viscosity at high concentrations

As regards the viscosity, SAKURADA[2] proposed an empirical formula which aims at enlarging the validity range of STAUDINGER's rule. According to this formula:

---

[1] W. GIBBS, *Principles of statistical mechanics.*
[2] I. SAKURADA *et al. J. Soc. Chem. Ind. Japan*, **37** (1934) 468, 470, 486, 487. Also: M. L. HUGGINS, *J. Am. Chem. Soc.*, **64** (1942) 2716.

$$\triangle \eta_{sp} = \frac{ac}{b-c} \tag{41}$$

where c is the concentration, and a and b are constants. It is best checked by plotting $c/\triangle \eta_{sp}$ against c, giving a straight line in the region where (41) applies. Schulz[1] uses the same expression, although in a different form. He writes

$$\frac{\triangle \eta_{sp}}{c} A(1 + B \triangle \eta_{sp}) \tag{42}$$

where A and B are constants. Solving for $\triangle \eta_{sp}$, this equation acquires the form of (41). According to Schultz, the constant B for a variety of macromolecules assumes values between 0.26 and 0.31.

Several authors[2] have, further, used a logarithmic viscosity formula:

$$\frac{1}{c} \ln \frac{\eta}{\eta_0} = \text{constant}.$$

In some cases this proves advantageous for extrapolation to zero c.

At still higher concentrations, formulae of the type (41) are no longer useful. A number of equations have been proposed by various authors[3], expressing the viscosity as a function of the concentration. Most of these equations try to cover a variety of colloidal solutions and are not restricted to long-chain molecules. For this reason they fall outside the scope of this chapter. Literature can be found in Philippoff's book[4] and in reference[3] (See further p. 121). We will here restrict ourselves to some remarks of a general nature.

b. Network in the solution

Some insight into the structure of concentrated solutions can be derived from the properties of randomly kinked molecules. It is clear that these molecules will get more and more entangled with one another as the concentration increases. This soon results in a more or less coherent sponge- or felt-like structure[5]. Accordingly, in these solutions no fractionation can be achieved by centrifugal means[6]: the felt migrates as a whole. Similarly, during diffusion a sharp boundary is maintained: the network of macromolecules remains untouched, while the solvent is as it were sucked into it. The process is therefore to be described rather as a swelling of the network than as a diffusion of the macromolecules.

c. Associative bonds

At some points of contact a temporary association may take place. These constitute the *knots* of the network mentioned ("Haftpunkte" in Frey-Wyssling's ter-

[1] G. V. Schulz and G. Sing, *J. prakt. Chem.* (2) 161 (1943) 161; E. Husemann and G. V. Schulz, *J. makromol. Chem.* (3) 1 (1944) 197.

[2] W. O. Baker, C. S. Fuller, J. H. Heiss, *J. Am. Chem. Soc.*, 63 (1941) 2142; G. Gee, *Brit. Rubber Prod. Res. Assoc. Publ.*, nr. 40; *Trans. Faraday Soc.*, 36 (1940) 1171.

[3] Literature can be found in H. L. Bredee and J. de Booys, *Kolloid-Z.*, 79 (1937) 31. See further H. L. Bredee and J. de Booys, *Kolloid-Z.* 79 (1937) 43; 91 (1940) 39; 99 (1942) 171; R. Houwink and K. H. Klaassens, *Kolloid-Z.*, 76 (1936) 217; 79 (1937) 138; 99 (1942) 160.

[4] W. Philippoff, *Viskosität der Kolloide*, Dresden-Leipzig 1942.

[5] J. Duclaux, *Rigidité, thixotropie, coacervation*, Paris 1934.

[6] K. H. Meyer and A. v. d. Wijk, *Kolloid-Z.*, 101 (1942) 53.

minology). In the liquid state these associations are continuously being formed and broken up again by BROWNian movement. The average lifetime of these links will depend on the nature of the solute and solvent and on the temperature. If either the temperature or the medium is changed to a sufficient degree, the links may become of a lasting character, thus leading to a gel; (compare Chapter on gels p. 483). This lasting character of the links in the gel is responsible for its elastic properties. Usually in the solution no elasticity exists[1]. As a result of their finite lifetime, however, the knots contribute to the viscosity of the solution[2], since the system is one of finite relaxation time and has elastic properties owing to the flexible macromolecules (compare MAXWELL's equation (22) on page 107).

d. Short-range order

It has been suggested by several authors[3,2] that the macromolecules in a concentrated solution are likely to show short-range order, which means that adjacent molecules tend to lie parallel to each other. The arguments advanced by STUART[3] to support this view are chiefly based on geometrical considerations and on model experiments with bead-strings in two dimensions. A still greater part in this phenomenon is probably played by intermolecular forces. On the strength of this argument it is reasonable to assume that some of the knots of the irregular network in the solution of macromolecules are of the nature of crystallisation germs. According to KARGIN and STEPANOWA[4] this is borne out by the phenomena observed during coagulation of cellulose acetate from organic solvents. At the addition of water, alcohol or cyclohexane to these solvents, the osmotic pressure drops practically instantaneously, while the viscosity increases slowly, to reach its ten-fold value after about a week. This is attributed to the slow formation and growth of crystallisation germs. The larger these germs, the longer is their average life (relaxation time in equation (22) and therefore the larger their contribution to the viscosity. It is of particular interest in this connection that the addition of small lumps of the gel to the solution greatly accelerates the viscosity increase: the formation of germs is thereby furthered considerably, in much the same way as this is brought about by the addition of crystals in the crystallisation process of substances of low molecular weight.

The energy content, and therefore the lifetime of a link will depend to a large extent on the number of macromolecules which take part in it. This number will become larger if the concentration is increased, which explains the rapid increase of viscosity with increasing concentration[2].

e. Time-effects

It is clear that, once a network is formed, the disentanglement of the macromolecules is no simple matter, and may require a considerable time. It is thus conceivable that certain properties of the solution depend on its pre-history. GLÜCKMANN[5]

---

[1] A very small real elasticity was observed in gelatin solutions by F. MICHAUD, *Ann. phys.*, 19 (1923) 63.
[2] S. A. GLÜCKMANN, *Acta Physicochim. URSS*, 13 (1940) 379; P. H. HERMANS, J. J. HERMANS and D. VERMAAS, *Kolloid-Z.*, 105 (1943) 199; J. J. HERMANS, *Kolloid-Z.*, 106 (1944) 95.
[3] O. KRATKY, *Koll.-Z.*, 68 (1934) 347; 84 (1938) 149; P. H. HERMANS, *Koll.-Z.*, 83 (1938) 71; H. A. STUART, *Naturwiss.*, 31 (1943) 123; 96 (1941) 301.
[4] V. A. KARGIN and A. A. STEPANOWA, *Acta Physicochim. URSS*, 6 (1937) 183.
[5] S. A. GLÜCKMANN, *Acta Physicochim. URSS*, 13 (1940) 379.

found that the 3% solution obtained by dissolving freshly coagulated cellulosenitrate shows a viscosity which is considerably larger than that of a 3% solution obtained by evaporating a 0.3% solution [1]. Results of a similar nature are reported by SAKURADA [2] who observed that the viscosity of celluloseacetate in a mixture of chloroform and benzene is higher if the substance is dissolved directly in the mixure than if the benzene is added afterwards.

According to FRANKEL [3] the change in some properties of gelatin solutions brought about by a change in the temperature could sometimes be followed over a period as long as 75 hours. MATHIEU [4] in X-ray measurements observed that the crystallinity of a cellulose nitrate film depends on the concentration of the solution from which it is obtained by evaporisation. This crystallinity also depends on the age of the solution: it decreases with increasing age, reaching a final value in about ten days. In the drying process the film shows a tendency to reassume the crystalline order of the original dry material but never succeeds in reaching this order completely.

f. Non-Newtonian viscosity in concentrated solutions

The coherent structure of macromolecules in concentrated solutions has more than once [5] been called upon to explain the "structural viscosity" of these solutions (decreasing viscosity at increasing rate of shear). The argument is quite simple in principle. In the liquid at rest the continuous formation and loosening of links by the temperature movement gives rise to a certain equilibrium concentration of the links. If a velocity gradient is applied, this concentration is affected by the shear. On the one hand the links are continually broken down by the shear. On the other hand there is a certain shear-making of links, which is of the nature of an orthokinetic coagulation. To a first approximation both the breaking-up process and the shear-making will be proportional to the rate of shear D. Hence, at large values of D, where the influence of the temperature movement becomes negligible, the stationary concentration of links which will be arrived at in the velocity gradient is independent of D. In that case the force F per unit area in the streaming liquid, being proportional to the number of links in unit volume, becomes independent of the shear-rate, i.e., the viscosity contribution of the links is

$$ \eta' = \frac{F}{D} = \frac{k}{D} $$

where k is a constant. Actually it appears [5,6] that the viscosity of many systems with

[1] A similar experiment was performed by H. STAUDINGER, *Cellulosechem.*, 20 (1942) 14.

[2] I. SAKURADA, *Kolloid-Z.*, 61 (1932) 50; His results were confirmed by H. ERBRING, *Kolloid-Z.*, 108 (1944) 16.

[3] M. FRANKEL, *Kolloid-Z.*, 45 (1928) 355; similar effects in solution of methylcellulose were studied by E. HEYMANN, *Trans. Faraday Soc.*, 31 (1935) 846.

[4] J. DESMAROUX and M. MATHIEU, *C. R. Acad. Sci.*, 194 (1932) 2053; M. MATHIEU, *Actual. Scient. Industr.*, 317, Paris 1936.

[5] W. KUHN, *Z. physik. Chem.*, A 161 (1932) 1; C. F. GOODEVE and G. W. WHITFIELD, *Trans. Faraday Soc.*, 34 (1938) 511; C. F. GOODEVE, *Trans. Faraday Soc.*, 35 (1939) 342.

[6] E. P. BINGHAM, *Fluidity and Plasticity*, New York 1922, p. 219; H. FREUNDLICH and E. SCHALEK, *Z. physik. Chem.*, 108 (1924) 153; A. SZEGVARI, *Z. phys. Chem.* 108 (1924) 175; E. HATSCHEK and R. S. JANE, *Kolloid-Z.*, 40 (1926) 53; R. V. WILLIAMSON, G. D. PATTERSON, and J. K. HUNT, *Ind. Eng. Chem.*, 21 (1929) 1111; C. ROSSI, *Gazz. chim. Ital.*, 67 (1937) 751.

"structural viscosity" at high rates of shear is closely approximated by a formula of the type

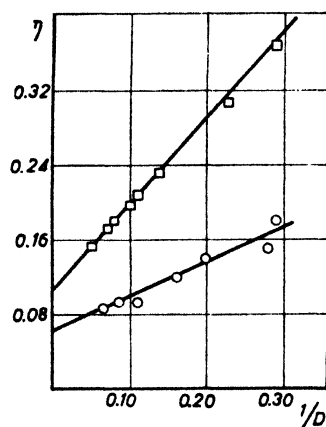$$\eta = \eta_0 + \frac{k}{D} \tag{43}$$

Fig. 11.
Viscosity versus shear-rate.
☐ „Baumwollgelb"
        0.42% in water
○ Sodium stearate
        0.1 % in water[1]

This is borne out by Fig. 11. For further particulars the reader is referred to GOODEVE's papers[2]. A full account on non-Newtonian viscosity can be found in PHILIPPOFF's[3] book.

g. Ultrasonics

The depolymerising action of ultrasonic waves, which has already been mentioned in Section 5 seems to be restricted to intermediate concentrations[4]. According to SCHMID[5] it is essential that the macromolecules should form a coherent structure, since presumably the isolated molecule is too small compared with the wave-length to be affected to any appreciable extent. On the other hand, at very high concentrations the structure becomes so strong that it is able to withstand the action of ultrasonics. This explains why the optimum effect is observed at intermediate con-
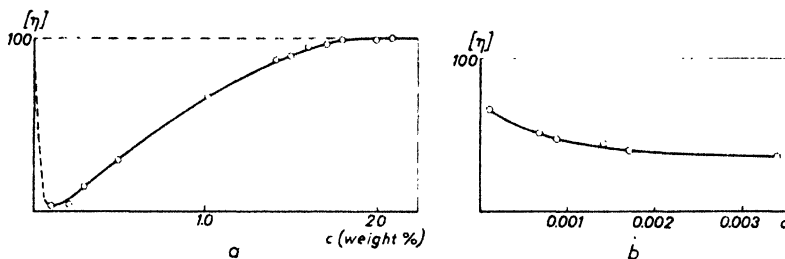
Fig. 12.   Depolymerising action of ultrasonics on caoutchouc in toluene[5].

centrations[6]. This maximum, however, is shifted towards lower concentrations as the size of the molecules increases, and is often not detected experimentally. Fig. 12 shows $\Delta\eta_{sp}/c$ in % of the original value for caoutchouc in toluene after irradiation. Since $\Delta\eta_{sp}/c$ is a measure for the degree of polymerisation, it is apparent that no de-

[1] H. FREUNDLICH and E. SCHALEK, Z. physik. Chem., 108 (1924) 175.
[2] W. KUHN, Z. physik. Chem., A 161 (1932) 1; C. F. GOODEVE and G. W. WHITFIELD, Tr.ıis. Faraday Soc., 34 (1938) 511; C. F. GOODEVE, Trans. Faraday Soc., 35 (1939) 342.
[3] W. PHILIPPOFF, Viskosität der Kolloide, 1942 p. 113 etc.
[4] Review article in H. MARK, J. Acoust. Soc. Am., 16 (1945) 183, and in P. GRABAR, J. chim. phys., 44 (1947) 145.
[5] G. SCHMID, Physikal. Z., 41 (1940) 326; G. SCHMID, E. BEUTTENMÜLLER, and A. RIEF, Kunst-stoff-Technik, 13 (1943) 65; G. SCHMID and E. BEUTTENMÜLLER, Z. Elektrochem., 49 (1943) 325.
[6] E. THIEME, Physikal. Z., 39 (1938) 384.

polymerisation takes place if $c > 1.8\%$. It is further seen that SCHMID's assumption according to which no depolymerisation occurs at very great dilution, although compatible with the experimental results, is by no means ascertained decisively. The dotted curve in Fig. 12 has been drawn on this assumption but in reality no measurements have been performed in this region.

§ 10. SOLID MATTER CONSISTING OF RANDOMLY KINKED MOLECULES

We will close this chapter by a discussion of the properties of *solid* matter consisting of randomly kinked molecules.

a. Elasticity of the dry polymer substance

It was shown by KUHN[1] that the elasticity of vulcanised rubber can be accounted for by the assumption of randomly coiled structures in the amorphous material. This material is, in principle, an irregular network of macromolecules[2]. The average molecular weight of these molecules is the average weight of the chain fragments connecting two successive knots[3] and is determined by the number of knots per unit volume. It need have no connection at all with the average molecular weight in the solution from which the gel originates.

If G is the number of knots (this is also the order of magnitude of the number of molecules) in unit volume, the modulus of elasticity is of the order of

$$E \sim kTG \tag{44}$$

This formula is essentially the same as equation (26) in Section 6, and may be derived by arguments which are identical to those given there. In fact, the contribution of a chain to the retractive force will be proportional to the elastic force kTr/NA^2 (equation 11, p. 99). The number of chain-fragments with length r, intersecting a cross-section of the material is proportional to Gr, if G represents the number of chains in unit volume. The total force is proportional to $kTGr^2/NA^2 = kTG$. The only difference[4] is that in Section 6a the argument was applied to free molecules dispersed in a solvent, while here it applies to the section of the chain between two successive junction points.

According to formula (44) the elasticity should be proportional to the absolute temperature. It was pointed out, however, by MÜLLER[5] that deviations from this rule need not necessarily indicate that the model does not apply. It is quite conceivable, that some of the knots which at one temperature have a lasting character, are of a more temporary nature at higher temperatures. They will then no longer

[1] W. KUHN, *Kolloid-Z.*, 68 (1934) 2; W. KUHN and F. GRÜN, *Kolloid-Z.*, 101 (1942) 248. Compare E. GUTH and H. MARK, *Monatsh.*, 65 (1935) 93; H. PELZER, *Monatsh.*, 71 (1938) 444; H. DOSTAL, *Monatsh.*, 71 (1938) 144, 309, 346; F. H. MÜLLER, *Kolloid-Z.*, 103 (1943) 144.
[2] A very elucidating support of the assumption of a network in elastic gels was given by W. F. BUSSE, *J. phys. Chem.*, 36 (1932) 2862.
[3] W. KUHN, *Z. Elektr.*, 45 (1938) 206; F. H. MÜLLER, *Kolloid-Z.*, 193 (1943) 144.
[4] The modulus of shear χ is closely related to the modulus of elasticity E; in incompressible isotropic media $E = 3\chi$. See, for instance, W. KUHN, *Z. phys. Chem.*, B 42 (1939) 1.
[5] F. H. MÜLLER, *Kolloid-Z.*, 103 (1943) 144.

play a part in the elasticity of the gel; in other words, the number G in equation (44) has decreased.

The relation (44) was derived by KUHN from entropy considerations. It was pointed out on p. 99, that a configurational entropy may be assigned to the macromolecule. When the molecules are deformed in the stretching process, they are forced into configurations of lower probablity, i.e., lower entropy. On this basis it is possible to calculate the free energy $F = U - TS$ of the stretched substance. If l represents the length of the material in the direction of stretch x, the retractive force K may be derived from F;

$$K = \partial F/\partial l = - T \, \partial S/\partial l \qquad (45)$$

since it is explicitly assumed (compare p. 94) that changes in molecular shape do not involve any change in internal energy U. To find the dependence of S on l, one must make some k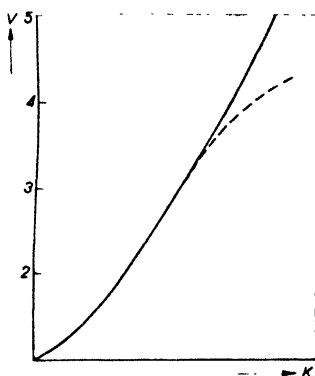ind of assumption regarding the mechanism of deformation. It is usually assumed that all chain-endpoints simply follow the changes in the macroscopic dimensions of the test piece. This means that if x,y,z are the coordinates of one endpoint with respect to the other one in the original unstretched material, the corresponding coordinates in the stretched substance are vx, $v^{-\frac{1}{2}} y$, $v^{-\frac{1}{2}} z$. Here v re-

Fig. 13A. Degree of stretch v plotted against force on cross-section of original substance.
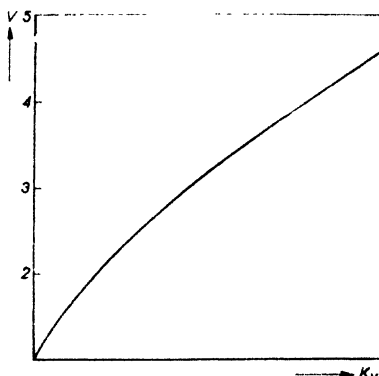
Fig. 13B. Stretch v plotted against tension, i.e., force per unit cross-section of stretched substance.

presents the ratio between the lengths before and after stretch and thus, since the volume remains practically constant, $v^{-\frac{1}{2}}$ is the ratio between the thickness after stretch and that in the unstretched condition. Accordingly, the distribution ψ_0 of chain-endpoints as given in equation (6) p. 98 is converted into

$$\psi = a' e^{- \dfrac{3}{2NA^2} \left(\dfrac{x^2}{v^2} + vy^2 + vz^2 \right)} \qquad (46)$$

From this new distribution the entropy of the stretched substance can be derived. In the limit of very small elongations (i.e., v close to unity) this leads to equation (44). For finite elongations one finds[1]

[1] F. T. WALL, J. chem. Phys., 10 (1942) 485; P. J. FLORY and J. REHNER, J. chem. Phys., 11 (1943) 521; L. R. G. TRELOAR, Trans. Faraday Soc., 39 (1943) 36, 242; W. KUHN and F. GRÜN J.Polym. Sci., 1 (1946) 196; J. J. HERMANS, Trans. Faraday Soc., 43 (1947) 591.

$$K = kTG \left(v - \frac{1}{v^2}\right) \tag{47}$$

This relation is represented in Fig. 13 A; it refers to the force per unit cross-section of the unstretched substance, and shows a curvature which is convex towards the force axis. In view of what follows, we anticipate at this stage that the K-v-relations found in practice are s-shaped (dotted curve in Fig. 13 A). This is due to the fact that the number N of chain-elements between successive junction points is finite (see below). Returning to equation (47), the *tension* in the stretched material, i.e., the force per unit cross-sectional area of the stretched substance is

$$Kv = kTG\left(v^2 - \frac{1}{v}\right) \tag{48}$$

and is plotted in Fig. 13B. Evidently, in this diagram the curvature is reversed. This was found experimentally[1] as early as 1921.

The distribution (46) of chain-endpoints was derived on the assumption that the endpoints simply follow the changes in macroscopic dimensions. A physical basis for this assumption is lacking[2]. The junction points in the network are not fixed in space and there is no obvious reason why they should follow the macroscopic deformation of the bulk substance. Rather must we expect that the endpoints will adjust themselves to some extent to the new situation and will occupy new equilibrium positions which are not necessarily those imposed by the macroscopic changes in the dimensions. An attempt to account for this can be made as follows. We assume that each chain-endpoint is, on the average, subject to a force p in the direction of the strain x. Introducing a potential energy $-px$, the new distribution of endpoints becomes

$$\psi' = a'' e^{-\frac{3r^2}{2NA^2} + \frac{px}{kT}} \tag{49}$$

Working this out[3], one arrives at a relation between force K and degree of stretch v, which is nearly the same as that given by Fig. 13A. From this point of view then, assumption (49) has no preference above (46). This second method of approach, however, proves to be advantageous in the treatment of more advanced degrees of stretch. In fact, the equations given so far assume tacitly that the number N of chain-elements between two successive junction points is very large. For all chain-fragments the ratio r/NA must be small compared with unity. At high degrees of stretch, this condition is no longer observed and it can easily be shown that, apart from corrections required in statistical considerations[4], the chain endpoints will no longer be able to follow the macroscopic deformation of the test-piece[5]. An extension of

[1] E. HATSCHEK, *J. Soc. Chem. Ind.*, 40 (1921) 251; R. ARIANO, *India Rubber J.*, 73 (1926) 271; R. HOUWINK, *Elast. Plast. Structur Materie*, Dresden and Leipzig 1938, p. 187.

[2] Compare, in this connection, GUTH and JAMES, *J. Chem. Phys.*, 11 (1943) 455, and the remark by J. J. HERMANS, *Trans. Faraday Soc.*, 43 (1947) 591.

[3] J. J. HERMANS, *Kolloid-Z.*, 103 (1943) 210.

[4] L. R. G. TRELOAR, *Proc. Phys. Soc.*, 55 (1943) 345; *Trans. Faraday Soc.*, 40 (1944) 109.

[5] J. J. HERMANS, *J. chim. phys.*, 44 (1947) 117.

the method based on equation (46) to finite N-values is therefore impracticable. Instead, the introduction of a potential energy $-px$ leads to a solution of the problem [1]. This theory makes use of KRAMERS' method mentioned in Section 8, p. 117, and leads to a force-strain relation which contains only two arbitrary constants. This relation is compared with experimental data for rubber in Fig. 14, which shows that agreement with theory is practically complete over the entire range of elongati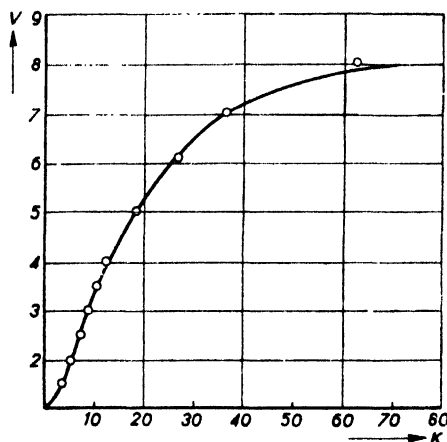ons attainable. A similar theory was developed by JAMES and GUTH [2]. TRELOAR's recent theory [3] of rubber elasticity for values of r/NA which are not negligible compared with unity is based on an extension of the distribution function η' to short chains. His results are also in conformity with experiment.

From what has been said in Sections 1 and 2 (p. 94) it will be clear that all *energy changes* on stretching have been neglected. The stress is due entirely to the tendency of the macromolecules to reassume their original shape, since the original molecule has *a greater probability* than the stretched one. It is obvious that this represents only an approximation [4] especially at high values of the stretch. In particular, the theory does not account for the effect of crystallisation at stretch (see p. 55).

Fig. 14. Comparison between theoretical and experimental force-strain relation (rubber).

However, the comparison of the theoretical curve with the experimental results in Fig. 14 suggests, that the influence of crystallisation and other energy effects on the stress-strain relation is comparatively small [5].

In connection with the stress-strain relations discussed here, it is interesting to observe that MOONEY [6] has shown how the retractive force in a stretching process is related to the shearing force in pure shear. We have seen on p. 107 that the shear modulus χ is defined by the ratio between shearing force and shear. If χ remains constant in finite deformation, i.e., if the shearing force is simply proportional to the shear, then the retractive force at finite stretch cannot be proportional to the elongation, and MOONEY finds that the retractive force is then related to the degree of stretch v as follows

$$K = \tfrac{1}{2}(1-v^{-3}) \; [(v+1) \chi + (v-1) \chi^*] \qquad (50)$$

provided the volume is not changed in the stretching process. Here χ^*, in addition to χ, is a constant characteristic of the material. The ratio χ^*/χ is called "coefficient of asymmetry", because it determines the difference between the behaviour at stretch and that at compression.

[1] J. J. HERMANS, *J. Colloid Sci.*, 1 (1946) 235.
[2] JAMES and GUTH, *J. Chem. Phys.*, 11 (1943) 455,
[3] L. R. G. TRELOAR, *Trans. Faraday Soc.*, 42 (1946) 83. 42 (1946) 83.
[4] Compare R. HOUWINK, *Z. Phys. Chem.*, A 183 (1932) 209.
[5] Compare also L. R. G. TRELOAR, *Trans. Faraday Soc.*, 38 (1942) 293.
[6] M. MOONEY, *J. Applied Phys.*, 11 (1940) 582.

All this is of particular importance to polymer substances of the rubber type, because both theory [1] and experiment [2] show that here the shear modulus is actually constant over a wide range of deformations. Accordingly, as TRELOAR [3] points out, the equation (47) is nothing but a special case of equation (5), to wit $\chi = \chi^* = kTG$..

b. Swelling of polymer networks

The theory of coiled structures in polymer networks has also been applied to swelling [4]. In many cases the cross-linking, which has formed the network, took place in the dry polymer. It is then reasonable to assume that in this dry condition the "normal" distribution (equation 6, p. 98) of chain-endpoints prevails:

$$\Psi'_0 = ae^{-\frac{3r^2}{2NA^2}} \tag{51}$$

Let us describe this situation as "normal coiling". In other cases cross-linking occurs in the swollen condition. Let us assume that coiling is normal at a degree of swelling q_0, this being the ratio between the volume of the swollen gel and that of the dry polymer. Then, at a degree of swelling q_1 coiling will be abnormal. If it is assumed that the chain-endpoints simply follow the macroscopic dimensions of the bulk substance, the distribution of endpoints becomes

$$\Psi = a'e^{-\frac{3r^2}{2NA^2}\left(\frac{q_0}{q_1}\right)^{2/3}} \tag{52}$$

From this distribution we can derive a formula for the entropy change at swelling:

$$\triangle S = kG \ln \frac{q_1}{q_0} - \frac{3}{2}kG\left[\left(\frac{q_1}{q_0}\right)^{2/3} - 1\right] \tag{53}$$

This is the change in configurational entropy. At the same time there is a change in the entropy of mixing. This change was considered in the chapter on thermodynamics (Section 6d, p. 72). We will not repeat the argument here, but simply assume that we have obtained an expression for the free energy of mixing in terms of the polymer-solvent ratio; i.e., in terms of the degree of swelling q_1. If $F_0(q_1)$ represents this free energy, the total free energy of the system may be found by adding the result (53):

$$F = F_0 - T\triangle S$$

If equilibrium is reached, the gel is in equilibrium with the pure solvent, which means that $\partial F/\partial n$ is equal to the partial GIBBS free energy of the pure solvent. This leads to a condition for the polymer-solvent ratio in equilibrium swelling.

Suppose now that the swollen gel is stretched. The result will be a further change in the distribution of chain-endpoints and therefore in the configurational entropy of the network. It can be shown that

---

[1] F. T. WALL, *J. Chem. Phys.*, 10 (1942) 485, cited after TRELOAR. This theory too assumes that the chain-endpoints follow the macroscopic deformation.

[2] Departures from linearity in pure shear were reported by L. R. G. TRELOAR, *Trans. Faraday Soc.*, 40 (1944) 59.

[3] L. R. G. TRELOAR, *Trans. Faraday Soc.*, 40 (1944) 59.

[4] P. J. FLORY and J. REHNER, *J. Chem. Phys.*, 11 (1943) 521; P. J. FLORY, *Chem. Rev.*, 35 (1944) 51; J. J. HERMANS, *Trans. Faraday Soc.*, 43 (1947) 591; 42B (1946) 155.

$$\triangle S = kG \left[ln \frac{q}{q_0} - \left(\frac{q_1}{q_0}\right)^{2/3} \left(\frac{v^2}{2} + \frac{q}{q_1 v}\right) + \frac{3}{2} \right]$$ (55)

where q is the degree of swelling after stretch, and v the ratio between the length of the stretched gel and that of the original swollen gel. It follows from this result that the partial free energy of the solvent, being proportional to $\partial F / \partial q$, is *decreased* at stretch, which means that *the degree of swelling will increase*. This increase is actually observed in a number of cases, e.g., in swollen cellulose nitrate [1], in cellulose acetate [2] and in reswollen cellulose model filaments [3], and may be considered as typical for coiling.

The retractive force in the stretch of the swollen gel is found by differentiating $-T \triangle S$ with respect to the degree of stretch v:

$$K = kTG \left(\frac{q_1}{q_0}\right)^{2/3} \left(v - \frac{q}{q_1 v^2}\right)$$ (56)

It includes equation (47) as a special case ($q = q_1 = q_0$), and shows how the elasticity of the network depends on the degree of swelling in the unstretched condition.

c. Double refraction of stretched rubber

Finally, let us briefly discuss the double refraction observed in stretching. The origin of this double refraction is similar in principle to that of the birefringence of flow dealt with in Section 7. In the unstretched substance there is no preferential orientation of the macromolecules and, therefore, no double refraction. If a stress is applied, the molecules are stretched and oriented. The resulting double refraction may be calculated from equation (4) on page 97 if the statistical distribution of the vector r in the stretched material is inserted [4]. We shall not give a derivation of the formulae concerned, but shall mention only an interesting feature of the result obtained. Whether one inserts the distribution (46) or (49) for the distribution of chain-endpoints, the resulting double refraction is always proportional to the tension Kv in the material. This also applies to the stretch of swollen gels if one inserts the corresponding distribution of chain-endpoints. In the special case, where all chain-endpoints follow the macroscopic deformation, one finds accordingly

$$\triangle n = c \left(\frac{q_1}{q_0}\right)^{2/3} \left(v^2 - \frac{q}{q_1 v}\right)$$ (57)

for the double refraction of the stretched swollen gel. It is clear that the constant c will contain the anisotropy $a_1 - a_2$ of the chain-element. This quantity may thus be computed from the experimental data. Comparison with the anisotropy of the monomer, gives an approximate value of the number v of monomeric groups in a chain-element. Considering that the monomeric groups in the chain-element show on the average a rather capricious broken line arrangement, so that the anisotropy of one monomeric group partially cancels that of the others, v will be about twice

[1] H. R. KRUYT, D. VERMAAS, and P. H. HERMANS, *Kolloid-Z.*, 100 (1942) 111.
[2] O. KRATKY and P. PLATZEK, *Kolloid-Z.*, 88 (1939) 78.
[3] P. H. HERMANS and P. PLATZEK, *Kolloid-Z.*, 97 (1941) 329; P. H. HERMANS, *Cellulosechem.*, 19 (1941) 122.
[4] W. KUHN and F. GRÜN, *Kolloid-Z.*, 101 (1942) 248; J. J. HERMANS, *Kolloid-Z.*, 103 (1943) 210.

the ratio between $\alpha_1 - \alpha_2$ and the anisotropy of the monomer. The result obtained in this way tallies well with the value derived from viscosity determinations. By way of example some data have been collected in Table 6. The double refraction has been denoted by $\triangle n$; the number L of monomeric groups per ml is easily obtained from the density of the substance and the molecular weight of the monomeric group; the number G of molecules per ml has been calculated from the elasticity modulus E; the degree of polymerisation P is obviously equal to L/G.

TABLE 6

ESTIMATION OF G AND ν IN VARIOUS SUBSTANCES

Substance	poly-styrene	polyvinyl-chloride	caoutchouc
temperature	150°	60°	20°
mol. weight of monomeric group	104	62	68
density (g/ml)	1.0	1.4	0.9
number L of monom. groups per ml	$6 \cdot 10^{21}$	$14 \cdot 10^{21}$	$9 \cdot 10^{21}$
elasticity modulus E, observed (dyn/cm²)	10^7	$3.5 : 10^7$	$1.3 \cdot 10^7$
G, calculated	10^{20}	$4 \cdot 10^{20}$	$1.5 \cdot 10^{20}$
degree of polymerisation $P = L/G$	60	35	60
$\triangle n$ at 100% stretch, observed	$4.5 \cdot 10^{-3}$	$2.2 \cdot 10^{-3}$	$2.0 \cdot 10^{-3}$
anisotropy $\alpha_1 - \alpha_2$ of chain-element, calculated . . .	$130 \cdot 10^{-25}$	$17 \cdot 10^{-25}$	$40 \cdot 10^{-25}$
anisotropy of monomer, observed	$50 \cdot 10^{-25}$	$3 \cdot 10^{-25}$	$20 \cdot 10^{-25}$
approximate value of ν	6	∽ 10	4

However, when applying results of the type (57), it must be borne in mind that such equations apply to the amorphous constituent only, and must fail at advanced degrees of stretch where crystallization sets in[1], since the influence of crystallisation on the double refraction of rubber-like polymers is considerable[2]. In addition to the intrinsic double refraction of the gel frame, the swollen gel may show structural double refraction and adsorption double refraction. The structural, or textural[3], birefringence is caused by the difference between the refractive indices of polymer substance and solvent. It was estimated by WIENER[4] for rod-shaped particles and for platelets. WIENER's theory is not applicable to rods of molecular thickness, and the structual double refraction caused by oriented long-chain molecules has not yet been accessible to a reliable theoretical treatment.

The adsorption double refraction results from orientated adsorption of solvent molecules on the polymer frame. There are indications that this effect plays a part in the double refraction of certain systems[5]. However, a theoretical approach to this type of birefringence does not exist. Notwithstanding these difficulties, TRELOAR[6]

---

[1] W KUHN and F. GRÜN, *J. Polymer Sci.*, 1 (1946) 199.
[2] L. R. G. TRELOAR, *Trans. Farad. Soc.*, 37 (1941) 84.
[3] A. FREY-WYSSLING, *Structure of the Protoplasm*, Elsevier, Amsterdam 1948.
[4] O. WIENER, *Abhandl. Sächs. Ges. Wiss., Math. Phys. Klasse*, 32 (1913) 508.
[5] D. VERMAAS, *Z. Physik. Chem.* B 52 (1942) 131.
[6] L. R. G. TRELOAR, *Trans. Faraday Soc.*, 43 (1947) 284.

has been able to show experimentally that the double refraction of both dry and swollen rubber, when stretched conforms reasonably to the requirements of the theory at low degrees of strain.

Of recent years the theory of rubber elasticity has been extended to networks in which the number N of chain-elements between successive junction points is very small [1]. The theory concerned makes use of the method developed by KRAMERS (see § 8 of this chapter, p. 117), which does not assume large values of N. The results obtained for the stress-strain diagram, the double refraction and the change in volume on stretching are in remarkably good agreement with the behaviour of cellulose gels if one takes N equal to about 1 in freshly prepared gels and to about 2 in reswollen gels [2]. In some respects this theory can be considered as an extension of KRATKY's studies on the behaviour of networks [3].

We shall close this chapter by pointing out that HUGGINS [4], in recent work, has developed a general theory of rubber elasticity, which does not make use of the concept of randomly coiled molecules. For this reason it falls outside the scope of this chapter.

[1] J. J. HERMANS, *J. Colloid Sci.*, 1 (1946) 235.
[2] J. J. HERMANS, *Trans. Faraday Soc.*, 42B (1946) 160.
[3] O. KRATKY, *Kolloid-Z.*, 70 (1935) 14; 84 (1938) 149, 268.
    J. J. HERMANS, *J. chim. phys.*, 44 (1947) 117.
[4] M. L. HUGGINS, *J. Polym. Sci.*, 1 (1946) 1.

V. THE DETERMINATION OF THE MOLECULAR WEIGHT OF MACROMOLECULES

J. J. HERMANS AND P. H. HERMANS

Laboratory for inorganic and physical chemistry, University of Groningen.
Laboratory for Cellulose Research of the A.K.U. and affiliated Companies, Utrecht.

§ 1. INTRODUCTION

The earlier literature referring to the subject of molecular weight determinations in macromolecular substances is very extensive and contains a great many contradictions and polemics. A review of this literature up to about 1935 can be found in the monograph by ULMANN[1]. In the last decade, however, much progress has been made and, though not all difference of opinion has as yet dissappeared, it would seem that a number of methods are available at present, which, according to the current opinion of the majority of scientists, enable one to perform more or less accurate determinations of molecular weight in a great many cases.

Among the methods which have been developed to determine the molecular weight of micromolecules there exist a few which are also applicable to macromolecules. Others, however, are impracticable, especially the so-called indirect osmotic methods such as the cryoscopic one and that based on the depression of vapour tension. The effects to be measured are too small when dealing with macromolecular substances. A substance with a molecular weight $M = 20000$ — which is still a relatively small value in the macromolecular field — would yield a freezing point depression of 0.0047 degrees in a 0.5% aqueous solution[2]. On the other hand, the osmotic pressure of such a solution corresponds to a water column of 62 mm height and can, therefore, be measured quite accurately. In addition a number of methods exist which have no counterpart in the region of low molecular weight, (viscosity, birefringence of flow, light-scattering, sedimentation in the ultracentrifuge). In a few special cases chemical methods can be used.

A point of particular importance in the determination of molecular weight of macromolecular substances is the fact that these substances usually represent mixtures of polymer homologues. In practically all methods we can go no further than to determine an *average* molecular weight. If no fractionation has been carried out, this average value may often refer to a spectrum of widely varying molecular

---

[1] M. ULMANN, *Molekülgrössenbestimmungen bei hochmolekularen Naturstoffen*, Dresden and Leipzig 1936.

[2] Moreover, the application of indirect osmotic methods in macromolecular substances sometimes gives rise to anomalies which have not yet been cleared up. This particularly refers to the method of "isothermic distillation" recommended by ULMANN (see footnote 1). For a discussion of these anomalies see e.g. O. KRATKY and H. MARK, *Fortschritte der Chemie organischer Naturstoffe* 1, 255 (1938); F. KLAGES, *Kolloid-Z.*, 93 (1940) 19.

weights. As will be shown in the following sections, the average value may then be different according to the method of determination used.

The reason of this is very obvious. The osmotic method and the chemical method yield the *number* of molecules in unit weight; in the viscosity method, in the method of sedimentation velocity and in the birefringence of flow either the *weight* or the *length* of the particle represent the decisive factor. In the former instances the effects to be measured decrease and in the latter they increase with increasing M. In solutions of equal concentration by weight small molecules increase the osmotic effect, but they have relatively little influence on, say, viscosity, whereas great molecules, if they are of the chainlike (linear) type, will have a small osmotic pressure but will strongly affect viscosity[1]. Consequently, the average values found in the two cases will greatly depend on the distribution of the molecular weights occurring in the polymer-homologous mixture. In order to characterize a given macromolecular substance by its molecular weight, it will, therefore, as a rule be necessary to determine this distribution as well. The procedure of fractional precipitation applied to this end is referred to in section 8 of this chapter (p. 144) and in the chapter on thermodynamics, section 7 b (p. 79).

The distribution curve of molecular weights may widely vary with the kind of chemical reaction in which the macromolecular substances were formed and will also depend on the further history of the sample. These items were studied, among others, by SCHULZ[2] in connection with the theory of precipitation. We further refer to MARK[3], KUHN[4], FLORY[5], MONTROLL and SIMHA[6], TUCKET[7], and to the chapter on sols, p. 153.

We shall express the concentration c in grams per liter. Let c_i denote the concentration of the molecules with molecular weight M_i. This means that there are $c_i .10^{-3} M_i$ moles per ml; in other words, the number of molecules with molecular weight M_i in 1 ml amounts to

$$G_i = \frac{N}{1000} \frac{c_i}{M_i} \tag{1}$$

where N is AVOGADRO's number. We assume that the total concentration

$$c = \Sigma_i c_i \tag{2}$$

in grams per litre has been determined by direct analysis. In this chapter several methods will be reviewed; the results obtained with different methods will be compared in section 10.

[1] H. STAUDINGER, *Die hochmolekularen organischen Verbindungen*; *Kautschuk und Cellulose* Berlin 1932 (p. 64 and 169) was the first to direct attention to this point. Its quantitative treatment has then been given by E. O. KRAEMER and W. D. LANSING, *J. Am. Chem. Soc.*, 55 (1933) 4319 and W. KERN, *Ber.*, 68 (1935) 1439. Cf. also the work of G. V. SCHULZ, cited below.

[2] G. V. SCHULZ, *Z. physik. Chem.*, B 32 (1936) 27; B 41 (1939) 466; B 46 (1940) 137; B 47 (1940) 155. G. V. SCHULZ and A. DINGLINGER, *J. prakt. Chem.*. 158 (1941) 149; G. V. SCHULZ, *J. makromol. Chemie*, 1 (1944) 131.

[3] H. MARK, *Ber.* 62 (1929) 1103; *Trans. Faraday Soc.*, 36 (1940) 611.

[4] W. KUHN, *Ber.* 63 (1930) 1510; *Z. Physik. Chem.*, A 159 (1932) 368.

[5] P. J. FLORY, *J. Am. Chem. Soc.*, 58 (1936) 1877.

[6] MONTROLL and SIMHA, *J. Chem. Phys.*, 8 (1940) 721; SIMHA, *J. Applied Phys.*, 8 (1940) 721, 12 (1941) 569; MONTROLL, *J. Am. Chem.*, *Soc.*, 63 (1941) 1215.

[7] R. F. TUCKETT, *Trans. Faraday Soc.*, 41 (1945) 351; R. A. BLEASE and R. F. TUCKETT, *Trans. Faraday Soc.*, 37 (1941) 571.

§ 2. CHEMICAL METHODS

If the macromolecule contains a well-defined number of groups which can be determined by chemical means, it is obvious that such a determination may be used to obtain the molecular weight of the substance. For instance, if each molecule carries a carboxyl group at the end, the number of these groups per unit of volume may be measured by direct or by conductometric[1] titration. Obviously, this method determines the total number of molecules, i.e., according to equation (1), the quantity $\Sigma_i c_i / M_i$. The average value obtained in this way is, therefore

$$\overline{M_n} = \frac{c}{\Sigma_i c_i / M_i} = \frac{\Sigma_i G_i M_i}{\Sigma_i G_i} \tag{3}$$

Obviously M_n is the "number average", because G_i is the number of particles with molecular weight M_i.

It is clear that the chemical method is restricted to a special type of macromolecule. Moreover, with high degrees of polymerisation, the amount of chemicals, used in the titration per gram of macromolecular substance becomes very small, which means that accurate measurements will be difficult. For this reason the method is restricted to molecular weights below ca 2.10^5. By way of example, we mention the determination of carboxyl groups in the condensation product of ω-hydroxy-decanoic acid, $HO(CH_2COO)_{p-1} CH_2COOH$, where P denotes the degree of polymerisation[2]. The determination of carboxyl groups has also been applied successfully in cellulose chemistry[3]. In protein chemistry the determination of sulphur or that of primary amino groups may play a similar rôle.

The application of the chemical method requires, of course, a sure knowledge of the chemical nature of the substance concerned. A famous example of a chemical method which has given rise to much confusion, as a result of the unexpected chemical changes occurring during its application, was the method of complete methylation of polymeric carbohydrates introduced by W. N. HAWORTH and M. MACHEMER[4].

§ 3. OSMOTIC PRESSURE

The osmotic method has proved to be a very useful tool in the molecular weight determinations of a great many macromolecular substances[5]. In addition it is a relatively simple method.

According to the classical formula of VAN 'T HOFF, the osmotic pressure in a very dilute solution containing c grams per litre is

$$\pi = \frac{RT}{1000} \frac{c}{M} \tag{4}$$

[1] E. SCHMIDT and coworkers, Ber. 60 (1927) 503; 67 (1934) 2037; Naturwiss., 19, 376 (1931) 1006.

[2] KRAEMER and LANSING, J. Am. Chem. Soc., 55 (1933) 4319.

[3] E. HUSEMANN and O. H. WEBER, Naturwiss., 30 (1942) 280; J. prakt. Chem., 161 (1942) 1; W. WEHR, Kolloid-Z., 88 (1939) 207.

[4] Cf. K. HESS and F. NEUMANN, Ber., 70 (1937) 710.

[5] Cf. also the article by G. V. SCHULZ in W. RÖHRS, H. STAUDINGER, and R. VIEWEG, Fortschritte der Chemie, Physik und Technik der makromolekularen Stoffe, Vol. II, München-Berlin 1942.

Hence, if the quantity π/c, termed the "reduced osmotic pressure" is plotted against c, a horizontal line should result. In the macromolecular field, and particularly if chain molecules are concerned, deviations from this simple law are usually met with even at great dilution (See Chapter III section 8, p. 84).

According to OSTWALD[1] the osmotic pressure may be represented by the empirical equation

$$\pi = \frac{RT}{1000} \frac{c}{M} + bc^n \tag{5}$$

where b and n are constants. To arrive at the correct value of M it is recommended by OSTWALD to follow the method introduced by BECKMANN[2] i.e., to determine the osmotic pressure at various concentrations and then to extrapolate the (π/c)-c-curve to $c = 0$. The equation

$$\lim_{c=0} \frac{\pi}{c} = \frac{RT}{1000 M}$$

is used to calculate M.

(For practical purposes the formula

$$M = 3.32.10^4 \, (1 + 0.00367 \, t) \lim \frac{c}{\pi} \tag{6}$$

may be used, where p = osmotic pressure in cm water, c = concentration in g/l and t = temperature in degrees Celsius).

The same method was successfully applied by HUGGINS[3], GEE[4] and others. From the statistical considerations in Chapter III p. 76 and 85, one derives a value of 2 for the exponent n in equation (5), as is borne out by a great number of experiments. In other words

$$\pi = \frac{RT}{M} c + Bc^2 \tag{7}$$

and a (π/c)-c- plot gives straight lines whose intercept with the ordinate determines M. The constant B depends on the solvent, but is independent of the molecular weight of the solute. In a few cases the third order term becomes significant. This third order term is derived from theory (see p. 85) and leads to the formula:

$$\frac{\pi}{c} - \frac{RTd_o}{3 M_o d^3} c^2 = \frac{RT}{M} + Bc \tag{8}$$

where d_o is the density of the solvent, d that of the solute, while M_o is the molecular weight of the solvent. Thus, instead of plotting π/c versus c, one should rather plot $\pi/c - RTd_o{}^2/3M_o d^3$ versus c.

According to SCHULZ[5], however, the (π/c)-c-curve shows a pronounced down-

[1] Wo. OSTWALD, *Z. physik. Chem.*, A 159 (1932) 375.
[2] BECKMANN, *Z. physik. Chem.*, 2, (1888) 720
[3] M. L. HUGGINS, *J. Am. Chem. Soc.*, 64 (1942) 1712; *Ind. Eng. Chem.*, 35 (1943) 216.
[4] G. GEE, *Trans. Faraday Soc.*, 36 (1940) 1162, 1171.
[5] G. V. SCHULZ, *Z. physik. Chem.*, 158 (1932) 237; A 176 (1936) 317; A 177 (1936) 453; *J. prakt. Chem.* (2) 159 (1941) 130; 161 (1942) 147; G. V. SCHULZ and A. DINGLINGER, *ibid.*, 158 (1941) 136; E. HUSEMANN, E. PLÖTZE, and G. V. SCHULZ, *Naturwiss.*, 29 (1941) 305.

ward curvature in the region of high dilution (compare Fig. 1). The experimental evidence for this effect, however, is far from conclusive[4]. SCHULZ has proposed the formula

$$\pi = \frac{RT}{1000} \frac{c}{M} \frac{1}{1 - cs} \qquad (9)$$

where s represents an "effective" specific volume. In a number of globular proteins[3] the value of s proved to be constant throughout a considerable range of concentrations. In solutions of long-chain molecules, however, s decreases with increasing concentration. SCHULZ believes that in such cases s is of the nature of a "swelling volume", and he uses the FREUNDLICH-POSNJAK equation

$$\pi = ks^{-\theta}$$

to relate s to the osmotic pressure. The theoretical aspects of SCHULZ's extrapolation have been discussed in Chapter III p. 83, where it was shown that no physical significance can be attached to the method[4]. It has shown some practical value in a few cases, but it cannot be considered as being of general applicability and may often lead to serious errors in the results obtained.

Fig. 1. Values of π/c plotted against c; ○ cellulose nitrate in acetone[1]; △ cellulose methyl ether in water[2].

The values of M obtained by extrapolating the (π/c)-c-curves are, as a rule, independent of the solvent used. This was first shown by DOBRY in a number of cases[5], and serves as a check on the assumption that the solute is monomolecularly dispersed. A further check on this assumption is to measure the influence of the temperature on the osmotic pressure[6]. If VAN 'T HOFF's law applies, one finds $d \ln \pi/d \ln T = 1$. Large deviations from this relation may be indicative of association between solute molecules since these associations are strongly dependent upon the temperature.

Compared with micromolecules the macromolecules show an advantage in that there usually exists no difficulty in finding a suitable semipermeable membrane. The lower limit of M at which the usual membranes become permeable for the solute lies at about 20000. It is necessary therefore to eliminate fractions below this limit before determinations in polymer homologuous samples are performed.

A simple apparatus to determine the osmotic pressure of macromolecular substances has been described by DOBRY and by SCHULZ[7]. To meet the disadvantage, frequently encountered in macromolecular substances, that the establishment of the osmotic equilibrium is slow and may even take several days, a special time saving arrangement with a *stirring* device has been proposed by ALBERT and KRATKY[8].
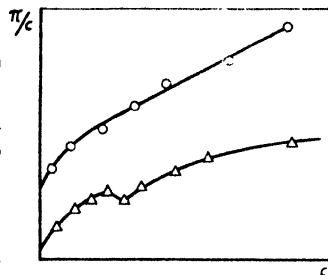
[1] G. V. SCHULZ, *J. prakt. Chem.*, 161 (1942) 151.
[2] G. V. SCHULZ, *Z. physik. Chem.*, A 177 (1936) 455.
[3] H. H. WEBER and R. STÖVER, *Biochem. Z.*, 259 (1933) 269.
[4] J. DUCLAUX, *J. chim. phys.*, 41 (1944) 209.
[5] A. DOBRY, *Kolloid-Z.*, 81 (1937) 90; *J. chim. phys.*, 32 (1935) 46.
[6] G. V. SCHULZ, *Z. physik. Chem.* A 180, (1937) 1; *Z. Elektrochem.*, 45 (1939) 652.
[7] G. V. SCHULZ, *Z. physik. Chem.*, 158 (1932) 237; *Z. physik. Chem.*, A 176 (1936), 317.
[8] O. ALBERT and O. KRATKY, *Oster. Chemiker Ztg.*, 1940, no. 7/8. See further R. O. HERZOG and H. M. SPURLIN, *Z. physik. Chem.*, A, Bodenstein-Festb., (1931) 239; P. J. FLORY, *J. Am. Chem. Soc.*, 65 (1943) 372.

To the same end a compensation method is sometimes used, in which the solvent is subject to a pressure of such magnitude that no osmosis takes place [1]. A still further technical simplification consists of applying a series of pressures and measuring the resulting velocity of osmosis. Extrapolating towards zero velocity, the osmotic pressure of the system at rest is obtained. The equilibrium method, however, is usually the more accurate one.[2] Recently, at SVEDBERG's suggestion, JULLANDER[3] replaced the direct measurement of the height of the liquid column by a gravimetric method, thus attaining an accuracy which, in principle, should be of the order of 100 times that of the length measurement.

Finally it is to be noted that incorrect results may be obtained if the macro-molecular substance exhibits a strong tendency to associate in the solvent chosen; this may occur even at extreme dilution. Such a case has been reported by STEURER[4] in solutions of cellulose ethyl ether in benzene[5]. In case of doubt more than one solvent should hence be used. Complications may further result from the presence of electrolytes. These give rise to DONNAN equilibria, which were treated in Volume I.

The osmotic pressure measures the number of molecules per unit volume. The average molecular weight is therefore the "number average" $\bar{M}_n$ (compare § 2). In fact, the contribution of c_i grams of molecular weight M_i to the osmotic pressure is

$$\pi_i = \frac{RT}{1000} \frac{c_i}{M_i}$$

The total pressure is $\Sigma_i\, \pi_i$, and thus the average molecular weight obtained conforms to equation (3)

§ 4. DIFFUSION

The diffusion coefficient of spherical particles in a liquid with viscosity η is given by the relation

$$D = \frac{kT}{6\pi\eta a},\tag{10}$$

where a denotes the radius of the particle; (compare Vol. I) From the value of a, determined in this way, the weight of the particle can be derived if the density is known. It is noteworthy in this connection that the volume, and therefore the weight, is proportional to a^3, which means that for instance an error of 5% in D gives rise to an error of 15% in a^3. With anisometric particles, formula (10) must be replaced by another expression. (See Vol. I). In which case the interpretation of diffusion data is less reliable.

As regards practical considerations, it is to be noted that the diffusion of colloid

[1] P. VAN CAMPEN, *Thesis*, Amsterdam 1930; *Rec. trav. chim.*, 50 (1931) 915; H. E. SLEUTEL, *Thesis*, Amsterdam 1936; CH. F. BOISSONNAS and K. H. MEYER, *Helv. Chim. Acta*, 20 (1937) 783; G. GEE, *Trans. Faraday Soc.*, 36 (1940) 1162.

[2] R. H. WAGNER, *Ind. Eng. Chem.*, 36 (1944) 520.

[3] I. JULLANDER, *Thesis*, UPPSALA 1945; *Ark. Kemi Miner. Geol.*, 21A, nr 8 (1945).

[4] E. STEURER, *Z. physik. Chem.*, A 190, 1 (1941) 16; *Kolloid-Z.*, 96 (1941) 333.

[5] By means of a small addition of a polar solvent (0.3% alcohol) the association could be repressed in this case (STEURER, *loc. cit.*).

particles is slow and therefore easily and considerably disturbed by convection currents or small temperature gradients. Moreover, the experimental methods are usually based on the assumption that the value of D is independent of the concentration, a condition which is not always met with in practice. Complications may further arise from electrostatic potential gradients if electrolytes are present (see Vol I). Notwithstanding these difficulties, the diffusion constant has been of great value in the study of proteins. Since these results can be found in SVEDBERG's book [1], we need not go any further into the matter.

The application of diffusion measurements to randomly kinked structures wants great caution. Here the change of D with increasing concentration is particularly large [2] since the molecules are soon interlinked. (See also the ensuing section). But even if it were possible to extrapolate safely towards zero concentration, the diffusion constant obtained would show no simple relation to the molecular weight. This is due to the hydrodynamic interaction of the chain elements, a phenomenon which has been dealt with in the chapter on kinky molecules, p. 103. According to the results obtained there, we must expect the diffusion constant of randomly kinked molecules to conform to an expression of the type

$$D \cong \frac{kT}{6\pi\eta Nr} (1 + | \overline{N}),$$

where N is the number of chain elements in the molecule and r the radius of a chain element. Experiments to verify this formula are not mentioned in the literature [3].

Recent work of JULLANDER [4] has thrown considerable doubt on the measurements of diffusion velocity in solutions of long-chain molecules. Further work in this field is very desirable.

§ 5.   ULTRACENTRIFUGE

The determination of molecular weight by ultracentrifugal means is of paramount importance in the study of proteins. The subject is treated in great detail in SVEDBERG's book [1], and we shall go no further than to discuss a few points of general interest. Two methods have been applied in this field, the determination of the sedimentation velocity and the study of the sedimentation equilibrium. In both methods we must distinguish between molecules with a randomly kinked loosely built structure and those which form compact particles, whether spherical or otherwise. This distinction must be made for practical reasons, because the theoretical interpretation of experimental results is much simpler and can be carried out on a safer basis in the case of compact particles than in that of loose structures.

If the particles are compact, their mutual interaction is small and only becomes important at comparatively high concentrations. Attempts to give a theoretical treatment of their hydrodynamical interaction in the sedimentation process were published by BURGERS [5]. Extrapolation to zero concentration is usually possible on

[1] THE SVEDBERG and K. O. PEDERSEN, *Die Ultrazentrifuge*, Leipzig 1940. See also Volume I.

[2] Compare, for instance, R. O. HERZOG and D. KRÜGER, *J. phys. Chem.*, 33 (1929) 179; *Kolloid-Z.*, 39 (1926) 250, 252; *Naturwiss.*, 14 (1926) 599.

[3] See, however, J. J. HERMANS, *J. Polymer Sci.*, 1 (1946) 233.

[4] I. JULLANDER, *Thesis*, Uppsala 1945, *Ark. Kemi Miner. Geol.*, 21 A (1945) 1 (English).

[5] J. M. BURGERS, *Proc. Acad. Amsterdam*, 44 (1941) 1045, 1177; 45 (1942) 9, 126.

a fairly safe basis. With randomly kinked structures, however, serious complications arise as a result of interlinking. Some of the aspects of this phenomenon have been discussed in the chapter on kinky molec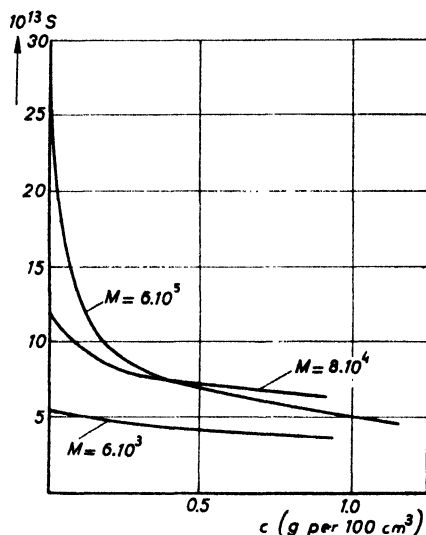ules, p. 119. Fig. 2 shows the extremely rapid rise of the sedimentation velocity with decreasing concentration in samples of cellulose nitrate[1]. In these cases the extra polation to zero concentration is liable to serious errors. We know further from section 4 in Chapter IV, p. 102, that the calculation of the molecular weight from these extrapolated values cannot be very accurate, since the hydrodynamic inter-action of the chain elements becomes involved.

Fig. 2. Sedimentation velocity, s, of cellulose nitrate in acetone as a function of concentration[1]

Turning our attention to the sedimentation *equilibrium*, we encounter similar difficulties. In principle it should be possible to derive the distribution of molecular weights in a mixture, from the concentration gradients in the sedimentation equilibrium. In practice, however, this is nos achievable. Even in polydisperse mixturet of compact particles a complete analysis along these lines is often impracticable. The more so with loosely built structures, where the molecules are too interlinked to sedimentate separately. The attainment of the equilibrium distribution, if at all possible, would require many weeks, or else such great dilution would have to be made use of, that experimental work would become practically impossible. This does not mean that all experiments in this field are illusory. However, some of the contradictions which still exist in the physical chemistry of long-chain molecules are doubtless due to difficulties encountered in experiments on their sedimentation (compare the examples in section 10 of this chapter). For an explanation of some of these contradictions see p. 103.

As regards the theoretical principles of the method, we will confine the discussion to ideal, i.e., to very dilute solutions. Considerations regarding the application of sedimentation to non ideal solutions can be found, among others, in papers by SCHULZ[2], GRALÉN[3], JULLANDER[4], LAMM[5].

The radial acceleration in a centrifugal cell, rotating with angular velocity ω, at distance x from the centre, is $\omega^2 x$. The centrifugal force per mole of solute with molecular weight M is, therefore $M\omega^2 x$. If the density of the liquid is ϱ, the centrifugal force results in a pressure gradient $\varrho\omega^2 x$, which exerts a force—$V\varrho\omega^2 x$ on the solute per gram of solute, if V is its specific volume. The total force per mole of

[1] H. MOSIMANN, *Helv. Chim. Acta*, 26 (1943) 61.
[2] G. V. SCHULZ, *Z. physik. Chem.*, 193 (1944) 168.
[3] N. GRALÉN, *Sedim. Diffusion Cellulose, Dissertation*, Uppsala 1944.
[4] I. JULLANDER, *Arch. Kemi Miner. Geol.*, 21A, nr 8 (1945).
[5] OLE LAMM, THE SVEDBERG 60*th anniversary volume*, Uppsala 1944, p. 182.

solute becomes $M (1 — V\varrho) \omega^2 x$. Introducing the frictional constant w per mole of solute, i.e., the friction experienced when moving with unit velocity, we obtain the following expression for the velocity of sedimentation:

$$u = \frac{M(1 — V\varrho)}{w} \omega^2 x \tag{11}$$

The frictional constant w is proportional to the viscosity η of the medium. In those cases where any of the quantities η, V, ϱ depends on the pressure to an appreciable extent, a correction must be applied to account for the influence of the pressure in centrifuging. For solutions in acetone this correction is estimated to about 30% in u if the centrifugal field is of the order of 4.10^5 times gravity [1].

As a result of the sedimentation, a concentration gradient is set up, and this gives rise to a diffusion which is superimposed on the sedimentation. The formula (11) assumes that the diffusion is not (or not yet) effective. The frictional constant w depends on M. With compact spherical particles w is proportional to the radius (STOKES' law) and therefore to $M^{1/3}$. We know from the chapter on random coiling, p. 103, that w in dilute solutions of kinky long-chain molecules is approximately proportional to $M (1 + \sqrt{N})^{-1}$, where N is the number of statistical chain-elements in the molecule. The quantity w can be eliminated, however, if we measure, in addition, the diffusion constant

$$D = \frac{RT}{w} \tag{12}$$

We find

$$M = \frac{RT}{D (1 — V\varrho)} \frac{u}{\omega^2 x} \tag{13}$$

This formula has been used by SVEDBERG and collaborators [2] to compute M in solutions of long-chain molecules. In this connection, however, we wish to refer once more to JULLANDER's work, which gives rise to some doubt on the experimental D-values (see the end of § 4, p. 137).

In a polymer homologous mixture, the M-value obtained from diffusion and sedimentation velocity shows no simple relation to number average or weight average. It can be shown, however, that it will usually lie between the two [3].

So much for the sedimentation velocity. If centrifuging is continued, an equilibrium distribution will ultimately be set up. The force $M(1 —V\varrho) \omega^2 x$ per mole of solute derives from a potential $— \frac{1}{2}M(1 — V\varrho)\omega^2 x^2$. Consequently, when equilibrium is reached, the concentration c at distance x from the centre of rotation is proportional to

$$c \backsim \exp. \left[\frac{M(1 — V\varrho)}{2 RT} \omega^2 x^2 \right] \tag{14}$$

In this state the sedimentation velocity is cancelled exactly by the velocity of diffusion. In practice one usually measures the gradient of the refractive index, which is pro-

[1] H. MOSIMANN and R. SIGNER, Helv. Chim. Acta, 27 (1944) 1123.
[2] R. SINGER, in THE SVEDBERG and K. O. PEDERSEN, The Ultracentrifuge; H. MOSIMANN, Helv. Chim. Acta, 26 (1943) 61; N. GRALÉN, loc. cit., I. JULLANDER, loc. cit.
[3] S. SINGER, Polym. Bull., 1 (1945) 79; I. JULLANDER, loc. cit.

portional to dc/dx, provided c is expressed in grams per unit of volume. Now, from (14) it follows that

$$\frac{1}{c}\frac{dc}{dx} = \frac{M(1 - V\varrho)}{RT} \, \omega^2 x \tag{15}$$

Thus, in a mixture of homologues, where the gradient of the refractive index is proportional to $\Sigma_i \, dc_i /dx$, one measures $\Sigma_i \, c_i \, M_i$, i.e., *the weight average molecular weight*. Strictly speaking, $\Sigma_i \, c_i \, M_i$ is not the same as $\Sigma c^\circ_i \, M_i$ in the liquid at rest, because the components with different M-values tend to produce different concentration gradients. In practice, however, a really effective separation of the components is not achieved and the difference is therefore of little importance, especially if the angular velocity is not too high.

§ 6. VISCOSITY

In a great many papers published since 1930, STAUDINGER has tried to establish that the specific viscosity increase due to long-chain molecules in dilute solutions is proportional to their molecular weight and conforms to the equation[1]

$$\triangle \eta_{sp}/c = B_\eta \, . \, M \tag{16}$$

where B_η is a constant which depends on the type of monomeric groups which constitute the macromolecule. This constant usually also depends (although in a lesser degree), on the solvent and on the temperature. STAUDINGER usually writes his equation in the form

$$\triangle \eta_{sp}/c = K_m \, . P \tag{17}$$

where P is the degree of polymerisation and K_m is another constant. STAUDINGER's relation was discussed in section 6 b of Chapter IV, p. 108, from the point of view of randomly kinked flexible structures. It was shown there that its applicability is restricted. Deviations from the equation are numerous, especially at high values of M and at very small values of M. Examples are given in the chapter mentioned.

To compute M, some authors recommend an extrapolation method. To that end the values of $\triangle \eta_{sp}/c$ are plotted against concentration and the curve is extrapolated to zero concentration[2]. An empirical formula of SAKURADA and SCHULZ was mentioned on p. 119.

The quantity $\lim_{c=0} \triangle \eta_{sp}/c$ is usually written $[\eta]$ and called ,,intrinsic viscosity''[3]. Its value depends on the definition of the concentration. If γ, γ_g and γ_v are the concentrations in g/cm³, g/g and cm³/cm³ respectively, ϱ_0 the density of the solvent and ϱ the density of the substance, we have

$$\triangle \eta_{sp} = [\eta]\gamma; \qquad \triangle \eta_{sp} = [\eta]_g \, \gamma_g \, ; \qquad \triangle \eta_{sp} = [\eta]_v \, \gamma_v \, .$$

The relation with STAUDINGER's constant in equation (17) is represented by

$$[\eta] = 10^3 M K_m \qquad [\eta]_g = 10^3 \varrho_0 M K_m \qquad [\eta]_v = 10^3 \varrho M K_m$$

---

[1] We adopt the notation which was introduced in Vol. I.
[2] H. STAUDINGER and W. HEUER, *Z. physik. Chem.*, A 171 (1934) 129; H. STAUDINGER and M. SORKIN, *Ber.*, 70 (1937) 1993; G. V. SCHULZ and F. BLASCHKE, *J. prakt. Chem.*, (2) 158 (1941) 130.
[3] Properly speaking one should add the condition $D \rightarrow$ o, where D represents the velocity gradient.

There have been many polemics as to the applicability of the STAUDINGER rule [1], which, however, need not be discussed here after what has already been said on the limits of the method. As regards the carrying out of viscosity measurements, the OSTWALD viscosimeter is a suitable apparatus. For the requirements as to the dimensions of the instrument we may refer to papers of BUNGENBERG DE JONG [2] and SCHULZ [3]. EKENSTAMM has given a simple formula to correct for the difference in density between solvent and solution [4].

In those cases where the relation (16) holds good, the average value of the molecular weight may be derived from viscosity measurement. This average value differs from that obtained in osmotic pressure experiments. The contribution of a molecule with molecular weight M_i to the specific viscosity of the solution is proportional to M_i. (Compare p. 111). The G_i molecules in unit volume with this molecular weight contribute an amount $G_i M_i^2$ which, according to equation (1) in the present chapter is proportional to $c_i M_i$. Consequently, the average molecular weight obtained in this way is

$$\overline{M_w} = \frac{\Sigma_i \, G_i \, M_i^2}{c} = \frac{\Sigma_i \, c_i \, M_i}{\Sigma_i \, c_i} \qquad (18)$$

This is the so-called "weight average"; each M_i is multiplied by the weight c_i of the substance with molecular weight M_i, then summed up, and divided by the total weight present. A very elucidating example was given by BAKER, FULLER and HEISS[5], who mixed a polymer with the corresponding monomer and found the viscosities given in Fig. 3. The straight line was calculated on the assumption that $[\eta]$ is proportional to $\overline{M}_w$, while the dotted curve refers to proportionality with the number average $\overline{M}_n$. The experimental values show conclusively that the viscosity is determined by the weight-average.

If equation (18) does not apply, it is, of course, possible to determine experimentally the relation between $\triangle \eta_{sp}$ and M in fractionated samples, where the values of M are restricted to a narrow region. Let Fig. 4 be the curve obtained. If now the

Fig. 3. Intrinsic viscosity of polymer-monomer mixtures.
o observed; — calculated from M_w;
- - - - calculated from M_n

unknown sample has also been fractionated, the value of M for a given value of $\triangle \eta_{sp}$ may be read from the curve. If, however, the unknown sample is a mixture of two homologues with molecular weights M_1 and M_2, the value of $\triangle \eta_{sp}/c$ is obtained by
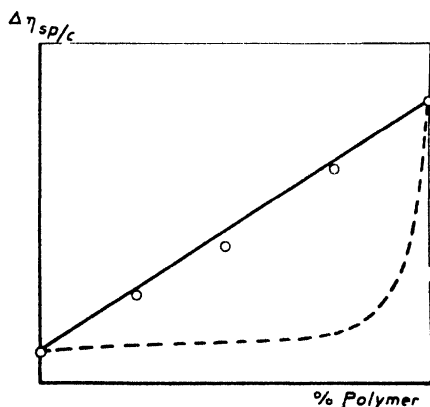
[1] See e.g., K. H. MEYER, *Kolloid-Z.*, 95 (1941) 70. Cf. also G. V. SCHULZ and A. DINGLINGER, *J. prakt. Chem.*, (2) 158 (1941) 163 and the answer of H. STAUDINGER, *Kolloid-Z.*, 98 (1942) 330.
[2] H. G. BUNGENBERG DE JONG, *First Report on Viscosity and Plasticity*, issued by *Acad. Sci.*, Amsterdam 1935, p. 110.
[3] G. V. SCHULZ, *Z. Elektrochem.*, 43 (1937) 479.
[4] A. AF EKENSTAMM, *Celluloselösungen in Mineralsäuren*, p. 82 (Lund 1936).
[5] W. O. BAKER, C. S. FULLER, J. H. HEISS, *J. Am. Chem. Soc.*, 63 (1941) 3316.

linear interpolation on the straight line AB. For instance, if the experimental value of $\triangle\eta_{sp}/c$ is given by the height C in Fig. 4, this means that the weight average of the molecular weight is M'. Yet the value derived from the curve is M''. Therefore, in this case the average molecular weight which is calculated from viscosity measurements in mixtures with an unknown distribution over the different M-values, shows no simple relation to the weight-average $\overline{M}_w$.

If (for fractionated samples) a relation of the type

$$[\eta] = KM^{\,n}$$

has been ascertained, the average M-value obtained from viscosity determinations is obviously [1]

fig. 4. Example of the relation between $\triangle\eta_{sp}$ and M found by experiment.

$$M_\eta = \left| \frac{\Sigma_i \, c_i \, M_i^{\,n}}{\Sigma_i \, c_i} \right|^{\frac{1}{n}} \qquad (19)$$

In spite of these uncertainties, viscosity determinations have been, and still are, of immense practical value in technical and scientific work. No scientific worker would easily forgo the great advantages furnished by the method, which is a quick and easy one, but he should exercise some caution in the interpretation of the experimental data.

§ 7. BIREFRINGENCE OF FLOW

Theoretical work in recent years [2] has thrown some light on the relation between the birefringence of flow in solutions of long-chain molecules and the molecular weight of these substances. The subject has been treated in section 7 of Chapter IV, p. 115. The contribution of a particle with molecular weight M_i to the double refraction $\triangle n$ is proportional to M_i [2] (see equation 31 on p. 116). Thus, for the value of $\triangle n/c$ we have an equation which is of the same form as (16), namely

$$\triangle n/c = B_n.M \qquad (20)$$

where B_n is a constant which is characteristic for the substance examined. As was shown in the preceding section, in a mixture of homologues the average molecular weight derived from $\triangle n$ will be the "weight average" $\overline{M}_w$. The application of equation (20) is subject to the same restrictions as that of (16). Thus, no ramifications or rings must occur in the molecule; the molecular weight must not be too large, and the randomly kinked structures must be loosely packed. Experiments are as yet too scarce to give judgment on the practical merits of the method.

A further quantity which is derived from measurements of double refraction

[1] J. W. TAMBLYN, D. R. MOREY, R. H. WAGNER, *Ind. Eng. Chem.*, 37 (1945) 573.
M. HUGGINS, *Ind. Eng. Chem.*, 35 (1943) 980; G. GEE, *Trans. Faraday Soc.*, 40 (1944) 261.
[2] W. KUHN and H. KUHN, *Helv. Chim. Acta*, 26 (1943) 1394; J. J. HERMANS, *Physica*, 10 (1943) 777; *Rec. trav. chim.*, 63 (1944) 25, 205; H. A. KRAMERS, *Physica*, 11 (1944) 1, *J. Chem. Phys.* 14 (1946) 415.

is the extinction angle Φ. The physical meaning of this angle may be appreciated from Fig. 10 on p. 115. It was there stated that tg 2Φ in a fractionated sample is inversely proportional to M^2. The influence of fluctuations in the molecular weight can be formulated as follows. If G_i is the number of molecules with molecular weight M_i, the value of tg2Φ becomes

$$\text{tg}2\Phi = B_\phi \frac{\Sigma_i G_i M_i^2}{\Sigma_i G_i M_i^4} = B_\phi \frac{\Sigma_i c_i M_i}{\Sigma_i c_i M_i^3} \tag{21}$$

Here B_ϕ is a numerical constant which depends on the type of macromolecule studied. It follows from this relation that the average molecular weight derived from the extinction angle is given by the equation

$$\overline{M_\phi} = \left(\frac{\Sigma_i c_i M_i^3}{\Sigma_i c_i M_i} \right)^{\frac{1}{2}} \tag{22}$$

The influence of fluctuations in the molecular weight on the value of $\overline{M_\phi}$ is very large. By way of example consider a uniform substance with molecular weight M_1 to which molecules with molecular weight $4 M_1$ are added. It suffices to add 6% by weight ($c_2 = 0.06 c_1$) to raise $\overline{M_\phi}$ to twice its value! This means that it will be difficult to check the theoretical relation between tg 2Φ and M^2, since sharp fractionation usually is no easy matter. Conversely, however, if the relation is reliable, it will be a valuable tool to investigate whether a fractionation has been succesful. For this reason the extinction angle promises to become a very useful quantity. Up to the present, however, the experiments mentioned in the literature are too scanty for us to enter into any details.

To conclude, we want to discuss a point of particular interest, to which attention was drawn by KUHN[1]. In the equations for $\triangle\eta_{sp}$, $\triangle n$ and tg2Φ the numerical constants B_η, B_n and B_ϕ depend on the substance examined. It was shown in Chapter IV that these constants depend on the number of monomeric groups in a chain element and on the effective length of these chain elements. Moreover, B_n is proportional to the anisotropy $a_1 - a_2$ of the chain element. However, from equations (28) p. 108 and (34) p. 116, it is seen that $\triangle\eta_{sp}$ tg 2Φ for a sharply fractionated sample is independent of the substance used. In fact, in both these formulae the quantity $N^2 A^3$ occurs, where N is the number of chain elements in the molecule and A the length of these chain elements. It follows that

$$\triangle\eta_{sp} \text{ tg } 2\Phi = K \frac{kT}{\eta_o D} G \tag{23}$$

Here K is a numerical constant which is *independent of the macromolecular substance*, η_o is the viscosity of the solvent and D the velocity gradient applied. Since the number G of molecules in unit volume is proportional to cM, we may also write

$$\frac{\triangle\eta_{sp}}{c} \text{ tg } 2\Phi = K \frac{kT}{\eta_o D} M \tag{24}$$

Thus, a mere determination of $\triangle\eta_{sp}$ and tg 2Φ at a given concentration c will suffice to enable the calculation of the molecular weight of the long-chain molecules, provided it is reasonably established that these molecules conform to the theoretical

[1] W. KUHN and H. KUHN, *Helv. Chim. Acta*, 26 (1943) 1394.

formulae mentioned and that the molecular weight is sufficiently uniform to make negligible the influence of fluctuations in M.

§ 8.   PRECIPITATION

If a non-solvent is added to a solution of macromolecules, these molecules are precipitated according to their molecular weight. This precipitation is indicated by the appearance of a turbidity in the solution. The theory underlying this method was given in the chapter on thermodynamics, p. 79. It was shown there that the method may be used to determine the distribution of molecular weights in a mixture of homologues ("fractional precipitation"). The largest molecules precipitate first, the smaller ones only at higher concentrations of the non-solvent. Since the solubility is usually strongly dependent on the temperature, the experiments must be carried out in a thermostat.

The concentration γ^* of the non-solvent at the moment when precipitation begins is usually related to the degree of polymerisation according to a formula of the following type:

$$\gamma^* = a + b/P^n \tag{25}$$

where a, b and n are constants. With loose structures of linear molecules the exponent n is practically 1. This applies for instance to cellulose nitrate in acetone, precipitated by water[1], to Polyvinyl acetate in toluene, precipitated by a methanol-water mixture[2], and to the methylesters of polymethacrylic acid in benzene, precipitated by cyclohexane[3]. With compact spherical particles we must expect n-values in the neighbourhood of 2/3, since in this case only the outer surface of the particles is subject to the action of the medium. Examples can be found in some proteins. Finally, if the long-chain molecule shows a pronounced ramification or if the randomly kinked structure is comparatively close-packed, n may assume values between 0.7 and 1. This is shown for instance, by branched polystryrene[1], by acetyl-starch[1] and by glycogene[4]. An accurate check on the value of n, however, is usually impossible. HUSEMANN's[4] experiments with glycogene, for instance, can be equally well described by $n = 2/3$ as by $n = \frac{1}{2}$ (Fig. 5). This is due to the fact that the value of P in this method is not unlimited. In practice the upper limit of the molecular weight lies in the neighbourhood of 5.10^5, the lower limit lies in the neighbourhood of 10^4. It is further to be noted that the constant b in equation

Fig. 5.   Dependence of γ^* on the degree of polymerisation for glycogene in water, precipitated by methylalcohol.

$$I = f(1/P) = 10^2/P;$$
$$II = f(1/P) = 1/P^{2/3};$$
$$III = f(1/P) = 10/P^{\frac{1}{2}}.$$

[1] G. V. SCHULZ and B. JIRGENSONS, Z. physik. Chem., B 46 (1940) 105.
[2] R. A. BLEASE and R. F. TUCKETT, Trans. Faraday Soc., 37 (1941) 571.
[3] G. V. SCHULZ and A. DINGLINGER, J. prakt. Chem., 158 (1941) 136.
[4] E. HUSEMANN, J. prakt. Chem., 158 (1941) 163.

(25) changes with the concentration. A series of experiments should therefore always be made at a given concentration.

Fractional precipitation is of very great practical value and its range of applicability is still increasing rapidly. In many instances it represents the only method by which the distribution of molecular weights in a sample may be determined. In principle this could also be achieved by ultracentrifugal means. We have seen, however, in section 5, p. 137 that this method is often hampered by serious practical obstacles.

It should finally be noted that the fractionation of polymer homologous mixtures usually gives the best results if the precipitate has the nature of a liquid, i.e., if a *coacervate* (see Chapter VIII p. 232) is precipitated[1]. An example is given by solutions of cellulose nitrate in acetone precipitated by the addition of water.

§ 9. OPTICAL METHODS

The birefringence of flow in solutions of long-chain molecules was discussed in Section 7. Other optical methods for the determination of particle size were mentioned in Volume I. One of them is the direct photographic observation in the *electron microscope*. By way of example we may mention the experiments with the rod-shaped molecules of tobacco-mosaic-virus[2]. Attempts to apply this method to the spherical molecules of glycogene, a polysacharide, were made by HUSEMANN and RUSKA[3].

The linear macromolecules forming loosely built structures are not accessible to this method, because only one of their linear dimensions lies within the range of electron microscopic visibility. Strongly ramified molecules, such as starch, however, although not really compact, give rise to pictures in the electron microscope which show no very distinct boundaries but which none the less allow of a rough determination of particle size.

Among the other optical methods we must mention the *small angle scattering* of X-rays, which in certain favourable cases may be used to determine the particle size (compare KRATKY's experiment, Vol. I).

STAUDINGER and HAENEL-IMMENDÖRFER[4] published a description of an interesting application of the TYNDALL *scattering* (see Vol. I) to the determination of particle size in glycogene solutions. Since glycogene molecules are very strongly branched, their volume is proportional to the molecular weight M. Consequently, the intensity of the light scattered at a given concentration by weight is also proportional to M. This was confirmed experimentally for M-values between 4.10^5 and 2.10^6; below values of 4.10^5 the method became unreliable; beyond 2.10^6 no osmotic check was possible. It will probably be possible, however, to extend these measurements to much higher molecular weights, since RAYLEIGH's law usually holds good up to diameters of about 10^{-5} cm[5], which for glycogene would correspond to $M = 10^9$.

The scattering of light represents a very general method of determining the molecular weight of large molecules, and much attention has been drawn to this

---

[1] In this connection we may refer to the important work of A. DOBRY, *J. chim. phys.*, 35 (1938,) 387; 36 (1939), 9.

[2] G. A. KAUSCHE, E. PFANKUCH, H. RUSKA, *Naturwiss.*, 27 (1939), 292.

[3] E. HUSEMANN and H. RUSKA, *J. prakt. Chemie*, 156 (1940), 1.

[4] H. STAUDINGER and I. HAENEL-IMMENDÖRFER; *J. makromol. Chem.*, (3) 1 (1944) 185.

[5] W. MECKLENBURG, *Kolloid-Z.*, 16 (1915) 97.

method in recent years[1]. To explain the principles of this method, we observe that a completely uniform medium is optically empty, i.e., it scatters no light. In a liquid or gas, however, density fluctuations occur which result in local changes in the refractive index, and therefore in light-scattering. In a solution we have, in addition, fluctuations in the concentration. Since the refractive index depends upon the concentration of the solution, as a result of these fluctuations the solution is rendered turbid. The magnitude of this light-scattering will therefore depend upon the extent of the fluctuations, and this extent can be calculated from statistical mechanics.

Anticipating the simple theoretical treatment given below, we can easily understand that the molecular weight M of the solute plays a part in the phenomenon. From simple statistics it is well-known that if the normal number of solute molecules in a given volume is ϑ, the fluctuations in this number are of the order $\vartheta^{\frac{1}{2}}$. Now, the concentration by weight is proportional to $c = M\vartheta$, which means that the fluctuations in c are of the order $M\vartheta^{\frac{1}{2}} = (Mc)^{\frac{1}{2}}$. The fluctuations in refractive index are proportional to those in c, and thus, since the light-scattering is proportional to the square of the fluctuations in refractive index, this light-scattering will be proportional to Mc.

Let us now proceed to more detailed considerations. We know that if two states of equal energy and volume are separated in their entropy value by an amount $\triangle S$, the probabilities of these two states are in the ratio exp. $(\triangle S/k)$. This follows at once from BOLTZMANN's relation $S = k \ln W$ mentioned on p. 66 in the chapter on thermodynamics. Thus, if the increase in the entropy of the solution as a whole is $\triangle S$ when a small volume element changes from its equilibrium state to another state at constant energy and volume, then the probability of this change is proportional to $\exp(\triangle S/k)$. If the change is assumed to take place at constant temperature and pressure, we must replace this by $\exp(-\triangle G/kT)$, where $\triangle G$ represents the change in the GIBBS' free energy.

Let c be the concentration in grams per unit of volume. A small volume v will then, on the average, contain cv grams of solute. If the concentration in this small volume becomes $c + x$, the original cv grams of solute will occupy a volume

$$v' = \frac{c}{c + x} v = v - \frac{v}{c} x$$

Clearly, this state can be attained by transferring a volume vx/c of pure solvent from the particular volume concerned to the rest of the solution. To that end we must transfer vx/cV_0 moles of solvent, if V_0 represents the molar volume of the solvent. Now, if g_0 is the molar potential of the solvent, the increase in GIBB's free energy when vdx/V_0 moles of solvent are transferred from a volume in which the concentration is $c + x$ to another volume in which the concentration is c, is given by

$$[g_0(c) - g_0(c + x)] \frac{vdx}{cV_0} = -\frac{v}{cV_0} \frac{dg_0}{dc} x\,dx$$

The total increase in free energy required to bring about a change $\triangle c$ in the concentration of the volume v is

[1] P. DEBYE, *J. Applied Phys.*, 15 (1944) 338; P. M. DOTY, B. H. ZIMM and H. MARK, *J. chem. Phys.*, 13 (1945), 159; An admirable review article was given by B. H. ZIMM, R. S. STEIN and P. DOTY, *Polymer Bull.*, 1 (1945) 90.

$$\triangle G = -\frac{v}{cV_o}\frac{dg_o}{dc}\int_0^{\varDelta c} x\,dx = -\frac{v}{2c}\frac{1}{V_o}\frac{dg_o}{dc}(\triangle c)^2 \tag{26}$$

This can be expressed in terms of osmotic pressure π, since we know from equation (26) p. 59 that $- dg_o/V_o dc = d\pi/dc$. In fact, we can also obtain the formula relating $\triangle G$ to π by the following simple reasoning: the osmotic work done in transferring a volume $v.dx/c$ of solvent from the volume v to the rest of the solution is

$$[\pi\ (c + x) - \pi\ (c)]\ \frac{v}{c}\ dx = \frac{d\pi}{dc}\frac{v}{c}\ x\,dx\ ;$$

integrating from zero to $\triangle c$ this gives

$$\triangle G = \frac{v}{2c}\frac{d\pi}{dc}(\triangle c)^2 \tag{27}$$

This, then, is the change in the free energy of the total system when a fluctuation of magnitude $\triangle c$ takes place in a volume element v. The probability that such a fluctuation occurs is proportional to exp. $(-\triangle G/kT)$, so that the average value of $(\triangle c)^2$ is given by

$$\overline{(\triangle c)^2} = \frac{\int_{-\infty}^{\infty} d\triangle c\ (\triangle c)^2\ \exp.\ (-\triangle G/kT)}{\int_{-\infty}^{\infty} d\triangle c\ \exp.\ (-\triangle G/kT)} = \frac{kTc}{v\,d\pi/dc} \tag{28}$$

A change $\triangle c$ in the concentration gives rise to a change $\triangle n = \triangle c\ dn/dc$ in the refractive index n. The *amplitude* of the lightwave scattered by a volume v in which the refractive index differs by an amount $\triangle n$ from that of its surroundings, is proportional ot $v\triangle n$, and the contribution of this volume to the *intensity* of the scattered light becomes, therefore, proportional to $v^2\ (\triangle n)^2 = v^2(\triangle c)^2\ (dn/dc)^2$. Using (28) we find that the average value of this intensity is proportional to

$$I = vkTc\ \frac{(d\,n/d\,c)^2}{d\,\pi/d\,c} \tag{29}$$

Summing up over all volume elements simply means replacing v by the total volume V of the system. We have thus obtained a formula for the light scattering in terms of osmotic pressure π. In the limit of very small concentrations, where VAN 'T HOFF's law applies,

$$\frac{d\pi}{dc} = \frac{RT}{M}\ \text{and}\ I \infty cM\left(\frac{dn}{dc}\right)^2$$

This shows how light-scattering can be used to determine the molecular weight M of the solute. The change of refractive index n with concentration c is usually effectively constant throughout the concentration range examined, and, for a given polymer, independent of M. Therefore, in a polymer homologous mixture, the intensity of light scattering determines $\Sigma_i c_i M_i$, i.e., *the weight average molecular weight* M_w.

In practice the value of M must be obtained by means of an extrapolation method,

since in polymer solutions, $d\pi/dc$, and therefore also I, is strongly dependent on c. If the solution conforms to the theoretical requirements laid down in HUGGINS' and FLORY's theory (see p. 85), we find

$$\frac{d\pi}{dc} = \frac{RT}{M} + \Psi(c),$$

where $\Psi(c)$ for a given polymer is independent of the molecular weight. DEBYE has shown how this may be used, in principle, to determine the distribution in a polymer homologous mixture.[3]

So much for the fluctuations in the concentration. The contribution of the density fluctuations is significant in solutions of micromolecules[1], but can usually be neglected if M is large. Its magnitude can be calculated along similar lines, if we remember that the work to be done on a volume v in order to bring about a change in hydrostatic pressure $\triangle P$ is

$$\int_0^{\Delta P} P\frac{dv}{dP}\, dP = \tfrac{1}{2}\, v\beta\, (\triangle P)^2$$

where β is the compressibility of the solution. This determines the probability of a given pressure-fluctuation. The average value of $(\triangle P)^2$ becomes $kT/v\beta$. If the refractive index n is known as a function of the pressure, this determines the average contribution to light-scattering. We find, in addition to (29) a term

$$I' = v\,\frac{kT}{\beta}\left(\frac{dn}{dP}\right)^2 \tag{30}$$

or, in terms of density ϱ,

$$I' = vkT\, \beta\varrho^2 \left(\frac{dn}{d\varrho}\right)^2 \tag{31}$$

This is simply to be added to the contribution (29)

The theory assumes that the fluctuations in adjacent volume elements are not correlated. ORNSTEIN and ZERNIKE[2] showed that this is permissible provided the range of the molecular interaction is very small compared with the wave-length of the light scattered. In polymer solutions, where the molecules themselves are large, this does not always apply. In addition, with large molecules it is necessary to account for an interference effect which so far has been neglected. We allude here to the interference between light waves scattered by different atoms in the same molecule. In small molecules these atoms are all so close together, that the phase difference between the waves scattered by different parts of the molecule is negligible. In large molecules, however, this no longer holds good. The effect on light-scattering will be to give an assymmetric intensity distribution: the scattering in the forward direction (i.e., the continuation of the primary beam) is more intense than that in the backward direction. This is at once obvious from a simple consideration of the light waves scattered (Fig. 6): the waves scattered from the points A

[1] R. GANS, Z. Physik., 17 (1923) 370; Hand. Exp. Physik, 19, Leipzig 1928.
[2] L. S. ORNSTEIN and F. ZERNIKE, Physik. Z., 27 (1926) 261.
[3] See, however, H. C. BRINKMAN and J. J. HERMANS, J. chem. Phys. 1949.

and B of the molecule in the direction P show a greater phase difference than those scattered in the direction Q. Consequently, measurements of light-scattering in relation to the angle of deflection will give information of the size and shape of the dissolved particles. For randomly coiled molecules the calculation has been carried out by DEBYE.

Further information regarding the shape of the particles may be obtained from the depolarization of the scattered light. This is due partly to assymmetry in shape, and partly to optical anisotropy of the particles. For further particulars, however, we refer to the literature cited and to the text-books concerned [1]. (See further Vol. I).
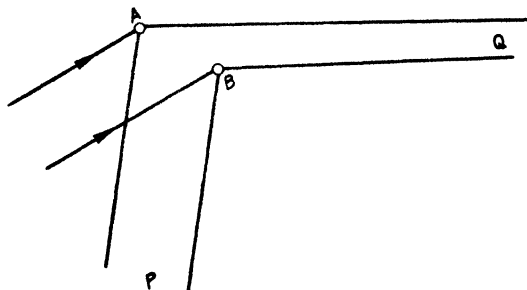
Fig. 6. Interference of lightwaves scattered in different directions (see text).

§ 10. COMPARISON OF RESULTS OBTAINED WITH DIFFERENT METHODS

The experimental check of one method by means of another often meets with considerable difficulties. This to a large extent is due to the difficulties encountered in the sharp fractionation of the samples. We have seen that different methods usually lead to different average molecular weights. If nothing is known about the distribution over the different M-values, a decisive judgment on the applicability of a method must usually be withheld. In this section some of the results known at present will be discussed briefly.

In the first place it is easy to show that the weight average dealt with in § § 6 and 7 (viscosity and double refraction), is always larger than the number average discussed in § § 2 and 3 (chemical method and osmotic pressure). To that end we must show that

$$\frac{\Sigma_i c_i M_i}{\Sigma_i c_i} > \frac{\Sigma_i c_i}{\Sigma_i c_i / M_i} \tag{32}$$

or

$$(\Sigma_i c_i M_i)(\Sigma_i c_i / M_i) > (\Sigma_i c_i)^2 ;$$

that is

$$\Sigma_i \Sigma_j c_i c_j M_i / M_j > \Sigma_i \Sigma_j c_i c_j .$$

Now, if in the left hand member the terms (ij) and (ji) are grouped together, we get $\frac{1}{2} \Sigma_i \Sigma_j c_i c_j (M_i / M_j + M_j / M_i)$. Thus we are to show that

$$\Sigma_i \Sigma_j c_i c_j \left(\frac{M_i}{M_j} + \frac{M_j}{M_i} - 2 \right) > 0,$$

[1] M. BORN, Optik, Berlin 1933; J. CABANNES, La diffusion moléculaire de la lumière, Paris 1929; R. GANS, Handb. Exp. Physik, 29, Leipzig 1928.

or

$$\Sigma_i \Sigma_j c_i c_j \frac{(M_i - M_j)^2}{M_i M_j} > 0.$$

This, however, is obvious, since all terms in this sum are positive. In a similar way it can be shown that the average value $\overline{M_\Phi}$ defined in equation (22) p. 143 is always larger than the weight average $\overline{M_w}$. The easiest way of doing so is first to derive the inequalities

$$\frac{\Sigma_i c_i M_i^3}{\Sigma_i c_i M_i^2} > \frac{\Sigma_i c_i M_i^2}{\Sigma_i c_i M_i} > \frac{\Sigma_i c_i M_i}{\Sigma_i c_i} \tag{33}$$

which can be done in the same way as that leading to (32). From (33) it follows immediately that

$$\frac{\Sigma_i c_i M_i^3}{\Sigma_i c_i M_i} = \frac{\Sigma_i c_i M_i^3}{\Sigma_i c_i M_i^2} \cdot \frac{\Sigma_i c_i M_i^2}{\Sigma_i c_i M_i} > \left(\frac{\Sigma_i c_i M_i}{\Sigma_i c_i}\right)^2;$$

q.e.d. It will now be clear what type of errors are to be expected when dealing with mixtures of homologues. For instance, if we plot $\triangle \eta_{sp}/c$ against the number average M_n obtained from osmotic pressure determinations, we derive a value of B_η in equation (16) which is too large. Many of the viscosity constants given in STAUDINGER's earlier work have thus been revoked in his later papers. Similarly, the value of tg 2Φ in experiments on the extinction angle will always be smaller than that found in a uniform sample with the same weight average. A reliable check can only be performed in sharply fractionated samples, or in samples where the distribution of molecular weight is known.

It was shown by HAWORTH and MACHEMER [1] that the molecular weight of cellulose determined by *chemical means* is confirmed roughly by *osmotic pressure* determinations. The method was improved by STAUDINGER and EDER [2], who found a quantitative agreement within a few per cent (M between 5 800 and 15 000). A similar result was obtained by them in the chemistry of polyhydroxy decanoic acids ($M = 8\ 000 - 39\ 000$). A further example of the chemical method checked by osmotic pressure determinations can be found for cellulose nitrate [3], ($M = 7 \cdot 10^4 - 35 \cdot 10^4$). Here, moreover, STAUDINGER's *viscosity rule* was found to apply. Further work on *viscosity* in conjunction with *osmotic pressure* was done, among others, by HUSEMANN and SCHULZ [4], who found that STAUDINGER's rule holds good for cellulose nitrate in acetone, ($M = 2 \cdot 10^4 - 4 \cdot 10^5$). If the samples were not fractionated, the value of $\triangle \eta_{sp}/c$ was too high, as would be expected. Less satisfactory results were obtained in solutions of polystyrenes [5] or polyesters [6], which is probably due to ramifications in these molecules. BAKER, FULLER, and HEISS [7] have shown that the corrected STAUDINGER re-

[1] W. N. HAWORTH and H. MACHEMER, *J. Chem. Soc.*, London (1932) 2270.
[2] H. STAUDINGER and K. EDER, *J. prakt. Chemie*, 159 (1941) 39; *Naturwiss.*, 29 (1941) 221.
[3] E. HUSEMANN and O. H. WEBER, *Naturwiss.*, 30 (1942) 115; *J. prakt. Chemie*, 161 (1942) 1.
[4] G. V. SCHULZ, *Angew. Chemie*, 49 (1936) 863; E. HUSEMANN and G. V. SCHULZ, *Z. physik. Chemie*, B 52 (1942) 1.
[5] H. STAUDINGER and G. V. SCHULZ, *Ber.*, 68 (1925) 2320; G. V. SCHULZ, *Z. physik. Chem.*, B 46 (1940) 137.
[6] H. STAUDINGER and H. SCHMIDT, *J. prakt. Chem.*, 155 (1940) 129.
[7] W. O. BAKER, C. S. FULLER and J. H. HEISS, *J. Am. Chem. Soc.*, 63 (1941) 2142.

TABLE

VARIOUS METHODS OF DETERMINING THE MOLECULAR WEIGHTS OF MACROMOLECULES

METHOD	APPLICABILITY RANGE	AVERAGE M IN MIXTURE	FURTHER REMARKS
chemical	$M < 2 \cdot 10^5$	M_n (eq. 3)	molecules with special chemical structure
osmotic	$M < 10^6$	M_n (eq. 3)	
diffusion	$M < 10^6$		not very conclusive; results as yet uncertain
sedimentation velocity alone	$M < 10^6$		not very conclusive in randomly coiled structures
diffusion + sedim. veloc.	$M < 10^6$	$M_n < M < M_w$ see text	still subject to uncertainties in randomly coiled molecules
ultracentrifuge equilibrium	$M < 5.10^6$	M_w (eq. 18)	in principle possible to determine complete distribution of M-values; in practice only in favourable cases
viscosity	unknown	M_w (eq. 18)	caution required, since STAUDINGER's rule is a limiting rule for loose structures
		M_η (eq. 19)	if M_n -law applies
birefringence of flow	unknown	M_w (eq. 18)	situation similar to that encountered in viscosity treatment
extinction angle	unkown	M_ϕ (eq. 22)	large influence of non-uniform molecular size
precipitation	$10^4 < M < 5 \cdot 10^5$		determines complete distribution of M-values
Light scattering	$M > 10^4$	M_w (eq. 18)	May also give informations on size and shape

lation, (see p. 111), applies to hydroxy-undecanoic polyesters, as checked by the *end-group method*, and the results of GEE[1] confirm the STAUDINGER law in relation to *osmotic measurements* in solutions of rubber in several solvents. In this latter case the range of M-values concerned is at least $6 \cdot 10^4 - 35 \cdot 10^4$, and probably greater. This result is of interest, because STAUDINGER's own measurements in rubber solutions

[1] G. GEE, *Trans. Faraday Soc.*, 36 (1940) 1171.

revealed considerable deviations from his viscosity rule, these being attributed to branching.

The agreement between *ultracentrifugal* data and that from *osmotic pressure* measurements is particularly good for solutions of globular proteins, for instance ovalbumin[1], haemoglobin[2] and serum albumin[3]. As regards linear macromolecules, not all difficulties are as yet overcome. In the determinations of *sedimentation velocity* and *viscosity* of solutions of cellulose in SCHWEITZER's reagent, and solutions of cellulose nitrate and cellulose acetate, the existence of a linear relationship between molecular weight and specific viscosity increase was confirmed[4].

A good agreement between *osmotic* and *ultracentrifugal* data exists also for solutions of cellulose methyl ether[5]. Later, however, a discrepancy was reported to exist in solutions of cellulose nitrate and cellulose acetate[6]. Some of these discrepancies could be explained[7] by the theoretical development laid down in the chapter on kinky molecules, p. 103. Moreover, we know from section 5 in the present chapter that ultracentrifugal work with randomly kinked structures is seriously hampered by the interlinking tendency even at comparatively high dilution. A good agreement between molecular weight determinations from *viscosity, birefringence of flow* and *ultra-centrifugal* data for solutions of cellulose methyl ether was recently reported by WISSLER[8].

As regards the *precipitation* method, the results obtained for solutions of the methyl-esters of polymethacrylic acid[9] were found to agree with *viscosity* and *osmotic* data, $(M = 15 \cdot 10^4 - 9 \cdot 10^5)$.

It is not our aim to give a complete summary of experimental results obtained so far, but the examples given here may suffice to give an idea of the present situation. We will conclude this chapter with a summary of the various methods discussed. The limiting values of the molecular weight M in this table are, of course, only approximate. They apply to long-chain molecules; a possible applicability range for sizes of compact particles is not considered here except in the case of light-scattering.

[1] N. F. BURK and D. M. GREENBERG, *J. Biol. Chem.*, 87 (1930) 197.

[2] G. S. ADAIR, *Proc. Roy. Soc. London*, A 120 (1928) 573.

[3] N. F. BURK, *J. Biol. Chem.*, 98 (1932) 353.

[4] E. O. KRAEMER, *Ind. Eng. Chem.*, 30 (1938) 1200.

[5] R. SIGNER and P. v. TAVEL, *Helv. Chim. Acta*, 21 (1936) 535.

[6] THE SVEDBERG and K. O. PEDERSEN, Leipzig 1940; see articles by KRAEMER and SIGNER.

[7] J. J. HERMANS, *Rec. trav. chim.*, 63 (1944) 219.

[8] A. WISSLER, *Thesis*, Bern 1941, cited after W. KUHN and H. KUHN, *Helv. Chim. Acta*, 26) (1943) 1394.

[9] G. V. SCHULZ and A. DINGLINGER, *J. prakt. Chem.*, 158 (1941) 136.

VI. MACROMOLECULAR SOLS WITHOUT ELECTROLYTE CHARACTER

R. Houwink

Rubberfoundation, Delft

---

§ 1. DISSOLUTION AND SWELLING OF HIGH-POLYMERS [1]

a. General considerations

Just as with the solubility of low-polymers, three different cases may be distinguished, namely:

1. Complete insolubility of the polymer in the solvent.

2. Solubility of the polymer and solvent in each other in certain proportions only.

3. Complete solubility of the polymer in the solvent, in all proportions.

Properly taken, 1 and 3 are the limiting cases of 2. In all three cases there is a partition equilibrium between the two phases; in case 1 the concentration of the polymer in the liquid phase becomes zero, in case 3 it can assume any value.

Case 2 offers the best starting point for tackling our problem. Just as with certain partly miscible low-polymer substances (e.g., phenol and water), restricted miscibility sometimes only exists below the so-called critical temperature T_{cr}; above T_{cr} complete miscibility (case 3) is observed.

A complication is often encountered with high-polymers, caused by swelling phenomena. In the older literature a very sharp distinction was made between dis-solution and swelling; the connection is however very close. For swelling, it is necessary that the penetration of solvent molecules into the solid is more rapid than the reverse process. For polymeric substances this condition is fulfilled; the small solvent molecules move more quickly than the large polymer molecules. Considered from this angle, swelling is thus a first step in dissolution [if no restrictions exist (see below)], since finally the slower macromolecules get the opportunity of dispersing themselves statistically.

A second question is whether the swelling (or dissolution) continues indefinitely in an excess of liquid, or whether it remains limited. In the latter case, only a well-defined volume increase of the polymer takes place and then the phenomenon of *limited swelling* is encountered. Here one is still dealing with a partition equilibrium of the polymer between the so called diluted part and the concentrated (i.e., gel) part. It will however appear that, when the molecular weight is sufficiently high, the concentration of the diluted part is so small that the conditions of case 1 are approached, and the material can be classified as practically insoluble.

---

[1] For a very comprehensive treatment of the subject from a purely physical point of view see J. H. Hildebrand, *Solubility of Non-Electrolytes*, New York, 1936. Also see p. 512 and 548.

Strictly speaking, this swelling can also be observed when mixing two low-molecular liquids. The fact that there is a partition equilibrium means that phase A contains a certain proportion of phase B and thus, phase A must be "swollen". If the volume of A before and after mixing [1] was determined (not a usual procedure), it should be possible to measure this partition equilibrium.

As has already been discussed above by BUNGENBERG DE JONG in Chap. I, in most cases, when dissolving macromolecular substances, sols of the *true thermodynamic equilibrium* type are formed. There are, however, also cases in which sols of the non-equilibrium type may be expected, namely when the macromolecules bear ionic groups and are dissolved in ionisable solvents.

Expressing this in terms of potential energy relations, treated by KRUYT in Vol. I of this book, both types of curves, namely that with an energy barrier and also that without energy barrier may be encountered.

The dissolution of high-polymers can be discussed not only from a thermodynamic but also from a molecular point of view.

b. Dissolution, considered from a thermodynamic stand-point

Starting from the example of two liquids, these will only mix as long as this involves a decrease of free energy. By the second GIBBS's law,

$$\triangle F = \triangle U - T \triangle S \tag{1}$$

$F = $ free energy, $U = $ internal energy
$S = $ entropy.

Hence liquids will mix when $\triangle F$ is negative, that means when $\triangle U$ is negative, or when $T \triangle S \rangle \triangle U$.

There are thus two factors influencing $\triangle F$, namely the internal energy and the entropy.

Whether $\triangle U$ will be positive or negative depends on various factors, connected with the attraction energies between the molecules.

The entropy however is always tending towards a maximum, leading to a positive $\triangle S$ and thus to a negative value of $\triangle F$. The process of dissolution is therefore always promoted by the entropy factor.

These general ideas, developed for the mixing of two liquids can also be applied to high-polymers, leading to the following considerations:

b. 1. The cohesion factor

For dissolution, two ground-molecules of the polymer must always be separated from each other and some solvent molecules as well, in order to create the conditions, necessary for a contact between these polymer and solvent molecules.

The energy $\triangle U$ involved is measured as the heat exchange on mixing and is expressed by:

$$\triangle U = -U_{LP} + (U_{LL} + U_{PP}) \tag{2}$$

where $U_{LP} = $ attraction energy between one ground-molecule of the polymer and its surrounding molecules of liquid (the so-called solvation energy).

---

[1] Taking into account that part of A is dissolved in B.

U_{LL} = attraction energy between the separated molecules of the liquid.

U_{PP} = attraction energy between two ground-molecules of the polymer.

When U_{LP} preponderates, energy is liberated (negative $\triangle U$) in the form of heat, and solvent molecules enter between the polymer molecules.

If however the term $(U_{LL} + U_{PP})$ preponderates over U_{LP}, then energy in the form of heat is to be added in order to let the solvent molecules penetrate into the polymer; without the addition of heat this would not be dissolved. This has appeared[1] to be the situation for many high-polymers.

b. 2. The entropy-factor

Even if there were no forces at all acting between the polymer and the solvent molecules, the polymer particles would still try to become scattered as regularly as possible all over the liquid, because the entropy aims at a maximum value. From the BOLTZMANN formula for the entropy:

$$S = k \ln W \qquad (3)$$

S = entropy
W = number of possible situations

it follows that W will be greater when the situations can also exist, that dispersed polymer molecules are scattered at random through the liquid, compared with the situation where this would not be the case. An example of this is already found in the mixing trend of two inert ideal gases. When these are brought into contact with each other without applying any external influence, they mix homogeneously. To give an idea of the order of magnitude[2] concerned in this entropy effect the average kinetic energy per molecule

$$E_{kin} = \tfrac{1}{2} kT \qquad (4)$$

and per degree of freedom may be indicated. At room temperature this is 10^{-13} to 10^{-14} erg per molecule and per degree of freedom. For the movement of the centre of gravity in the three dimensions of space this would give an amount of about 1000 cal per mol, reaching the order of magnitude of the secondary energy (see p. 156) per atomic group, exerted by the macromolecules on each other. However[3], for the case of macromolecules instead of micromolecules — as these are present in a mixture of two liquids — a considerably larger entropy of mixing per mol must be taken into account.

Thus it can be explained that even with a solvent causing no heat development, dissolution still occurs, due entirely to the entropy factor. By increasing the temperature, this factor becomes all the greater and so it can be understood that for certain combinations of polymers and solvents a critical temperature (T_{cr}, see p. 159) can be observed, above which there is miscibility in all proportions; in this case the complete cohesion factor is overcome by $T\triangle S$. Theoretically such a T_{cr} should

[1] G. GEE, Trans. I.R.I., 18 (1943) 266.
[2] Cf. H. MARK, The General Chemistry of High-polymeric Substances, Amsterdam, 1940, p. 213.
[3] P. J. FLORY, J. Chem. Phys., 10 (1942) 51, M. L. HUGGINS, J. Chem. Phys., 9 (1941) 440.

exist for any combination of polymer and solvent, practical limits however being given by the boiling point of the solvent or the distortion temperature of the polymer.

c. Dissolution considered from a molecular stand-point

When the bonds between the macromolecules are of the *homopolar* type, they cannot be loosened either by polar or by non-polar liquids at room temperature and this is the reason for the insolubility of substances, interlinked in this way, in any solvent at not too high a temperature. Phenolformaldehyde resin in its hardened state forms a striking example; it is insoluble in alcohol and similar solvents. It can be dissolved by boiling at very high temperatures (β-naphtol, b.p. 286°C), but for quite a different reason. This must be ascribed to the fact that so much thermal energy is then added that breaks in the macromolecules result, which in many cases are at places other than the interlinking bridges. In such cases destruction-products, differing from the monomer, are obtained.

If the primary bonds are of the *heteropolar* type, quite a different behaviour can be observed. These are not loosened by non-polar solvents, but are very easily loosened by certain polar solvents, just as NaCl dissolves easily in water as a consequence the very high hydration energy of the latter. Examples of this type have already been given in Fig. 8 p. 34 for some proteins bearing ionic groups. They will be treated in Chapter VII by OVERBEEK and BUNGENBERG DE JONG.

The coherence by means of the weaker *secondary bonds* shows some analogy with the heteropolar type, in so far that here also the bonds are loosened by certain groups of solvents, and not by others, depending on their chemical constitution (see p. 34)

We now come back to the phenomenon of *limited swelling*, already noted on p. 153 and which is often encountered in macromolecular substances to a very important degree, especially when chain-molecules are present, consisting of different kinds of monomer groups exerting a different cohesive energy. In this case, the possibility exists that the attachment between certain groups is loosened by the solvent, whereas other groups are still coherent to each other. In this way the solvent can penetrate freely at certain spots whereas at other places the coherence in the gel is maintained, leading to one of the pictures of Fig. 1.

With regard to the causes for the coherence at certain spots, four different cases can be distinguished, namely:

a. The presence of chemical bridges, interlinking the macromolecules by *primary bonds* $G_2 — G_2'$ in Fig. 1a.

b. The mutual attraction of certain groups on the macromolecules, leading to the formation of *secondary bonds*, $G_2 — G_2'$ in Fig. 1b, of a different chemical type from the secondary bonds between the other groups G_1. A condition for gel-formation is that there is a development of heat between the solvent and the groups G_1, or a consumption of heat with such a small value, however, that it can be compensated by the entropy effect. When at the same time so much energy is required for the loosening of the G_2 groups that it cannot be supplied by the entropy effect, only a local loosening of the chains takes place.

c. The case based on potential curves, discussed already by KRUYT in Volume I

a) As a consequence of primary forces (chemical interlinking).

b) As a consequence of secondary forces. (The groups G_2 are not loosened by the solvent, in contrast to the groups G_1).

Causes for limited dissolving (swelling)

c) As a consequence of secondary forces with a long range of action.

d) As a consequence of the great length of the macromolecules involved.
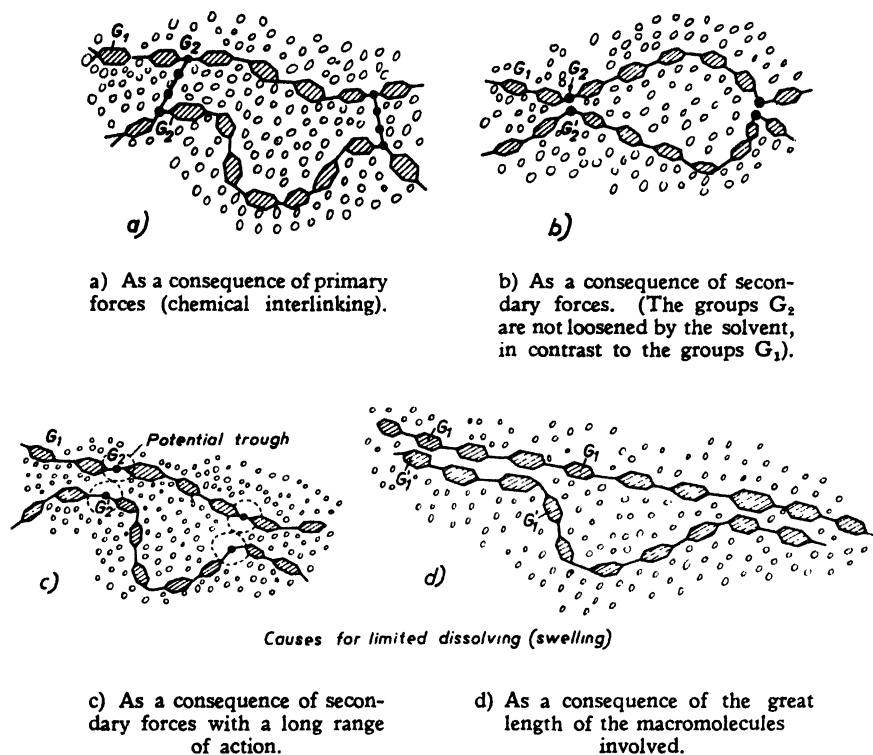
Fig. 1. Causes for limited dissolution (swelling).

of this book, showing a *minimum of potential energy at a certain distance from the particle surface*[1].

Although there is no proof available that this type actually exists in macromolecular substances, we will keep this possibility [2] open. Under these circumstances a gel would be obtained without any material contact between the macromolecules, leading to the picture of Fig. 1c. It may be that in macromolecular sols with electrolytic character, where ionic swarms cause special attractive and repulsive forces, such gels are encountered.

d. A fourth cause of gel formation is met with in Fig. 1 d, where the coherence by a series of G_1 groups only, as a consequence of the *great length of the macromolecules*, is pictured. This will be discussed on p. 160.

---

[1] HAMAKER applies this type of curve to "lyophilic" sols only. In our opinion this application is too limited. We will see, for instance, on p. 178, that certain sols with spherical macromolecules (phenol-formaldehyde resin in acetone) behave typically according to this type of curves, whilst they cannot be considered as "lyophilic", because solutions up to a concentration of 50% can be made, still having a low viscosity.

[2] The possibility also exists that such a long-range interaction exists for all groups. This case is not represented here.

§ 2.   THE INFLUENCE OF VARIOUS FACTORS ON THE SOLUBILITY

a.   The relation between the cohesion factor and the entropy factor

Only for some cases is the relation between the internal energy on mixing (cohesion factor) and the entropy factor known with certainty. For raw rubber, it is shown in Fig. 2 as a function of the swelling degree. The striking conclusion is that even for a solvent such as benzene, which is generally found in the literature to have a strong wetting power for the rubber molecules, the solubility is a consequence of the considerable entropy increase alone.
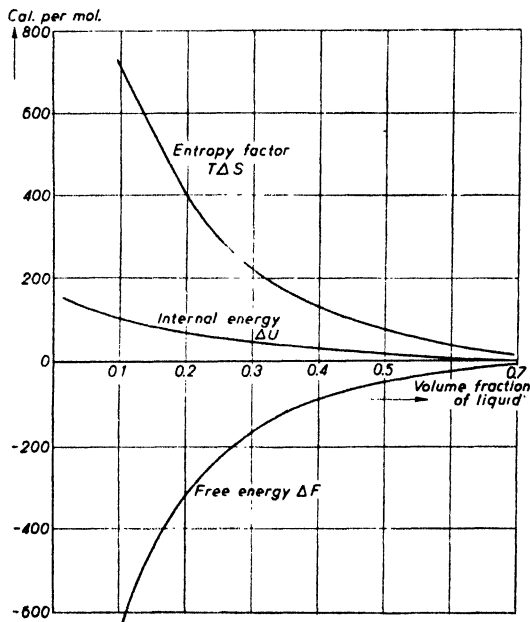
Fig. 2. Thermodynamic constants as a function of the swelling degree for rubber in benzene.
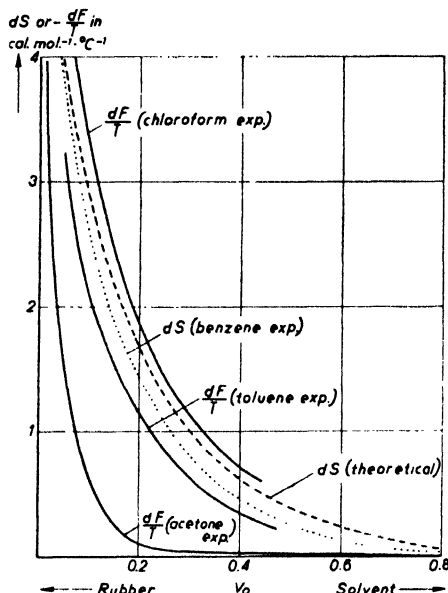
Fig. 3.   Entropy- and free energy changes when dissolving rubber in various solvents.

The internal energy, measuring the cohesion factor is positive [1] and so *rubber would not swell at all in benzene, if the entropy factor were not playing a part.*

Since the entropy is independent of the kind of liquid [2], the different swelling action of the various solvents must be ascribed to their difference in $\triangle U$ with regard to rubber; for good solvents the $\dfrac{\triangle F}{T}$ curve is on a lower level than for poor solvents, as is shown in Fig. 3. In this picture it is shown that the theoretically calculated values of dU are in satisfactory agreement with experimental values. Secondly it appears that $-\dfrac{\triangle F}{T}$, in the excellent solvent chloroform, is appreciably larger than in the poor solvent acetone, whereas toluene has an intermediate position.

---

[1] This means that heat has to be added.
[2] G. GEE, *Trans. Faraday Soc.*, 38 (1942) 418.

b. The influence of temperature

It can be calculated [1] that at a sufficiently high temperature, the change of the GIBBS's free energy of dilution with concentration $\left(\dfrac{\partial F}{\delta Cpol}\right)T$ is negative for all values of $Cpol$. Above this critical temperature T_{cr} the polymer and the liquid will be miscible in all proportions. Below T_{cr} two phases will coexist and the solubility is highly temperature dependent. That most liquids are either non-solvents or completely miscible with a polymer can be ascribed to this last fact, the temperature range in which the solubility is limited and measurable being confined to a few degrees only. Below T_{cr} the dilute phase consists essentially of pure solvent, and the solid phase may contain appreciable quantities of solvent.

Fig. 4 shows some data for rubber in benzene, calculated by GEE. Below, we return to the influence of the molecular weight, referring now only to the very pronounced temperature influence. For rubber with a mol. weight of 10^5, the temperature range of solubility is only from about 0 to 25°C; at 0°C the solubility becomes practically unmeasurable, at 25°C complete miscibility is reached.

There are cases where T_{cr} is higher than the boiling point of the solvent, for example in the case of cellulose and water. Then no complete dissolution can be obtained by heating, and chemical means must be introduced. Either the composition of the solvent must be adapted — a procedure much encountered in paint- and varnish-industry — or the constitution of the polymer has to be changed, a method followed in the rayon industry (cellulose $\longrightarrow$ cellulose xanthogenate). Such problems will be reconsidered in this book by P. H. HERMANS.
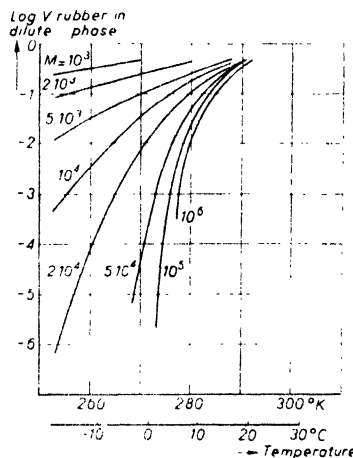
Fig. 4. Effect of temperature on the solubility of rubber in benzene for various molecular weights.

c. The influence of the polymerisation degree

Experimental evidence is available, showing that, in increasing their size, soluble macromolecules may finally become insoluble.

This can be formulated mathematically on a basis of studies by BRØNSTED [2], from which it became accepted that when dissolving a polymer, the concentration decreases proportionally to e^M. Further considerations lead to the relation:

$$\frac{C_p}{C_P} = e^{(P-p)q/RT} \tag{5}$$

[1] G. GEE, Trans. Faraday Soc., 38 (1942) 276.

[2] J. N. BRØNSTED, Z. phys. Chem. Bodenstein, Festband (1931) 257; C. r. Lab. Carlsberg (chimie), 22 (1938) 99. For criticism see: G. GEE, Advances in Colloid Science, New York, 1946, p. 175.

where C_p and C_P are the concentrations in the solution of two fractions with polymerisation degree p and P respectively.

q = difference of potential energy per structural unit.

This means that the solubility is strongly dependent on molecular weight and temperature. To illustrate this, it may be realized that, on increasing the degree of polymerisation from $p = 10$ to $P = 100$, the concentration C_P will become 10^{10} times smaller than C_p. In the case where the low-polymer p is already poorly soluble, this results in the fact that, by increasing the molecular weight only slightly, the solubility becomes immeasurably small, leading to the conclusion that the material is practically insoluble. It can be said that the macromolecules contract themselves completely into the undissolved part on polymerisation.

This influence of molecular weight on the solubility can be demonstrated by means of low molecular paraffins. The solubility diminishes by about 70% for every newly added CH_2—group, as appears from Table 1.

TABLE 1

SOLUBILITY OF NORMAL PARAFFINS

	solubility in water in mol/l($\times$ 10⁴)
C_5H_{12}	50
C_6H_{14}	16
C_7H_{16}	5.2
C_8H_{18}	1.3

A further example is found in the case of cellulose. Crystallised glucose (mol. wt 180) dissolves easily in water; cellobiose (mol. wt 342), cellotriose and cellotetrose dissolve with rather more difficulty whilst crystallised cellulose is completely insoluble. Amorphous cellulose has an intermediate position; it only swells in water, a point to which we will return on p. 165.

The figures of Table 2 correspond to Fig. 4; they show that the critical temperature varies only between 273 and 296°K for a range of molecular weights between 10^3 and 10^5. In contrast to this the change in solubility is from about 10^{-1} to 10^{-17} (vol. fraction in the dilute phase).

TABLE 2

CRITICAL TEMPERATURES AND COMPOSITIONS OF PHASES AT-10° C FOR RUBBER-BENZENE SOLUTIONS

Mol. wt	T_{cr} in °K	Vol. fraction of rubber	
		in the dilute phase	in the concentrated phase
1.10^5	296	1.10^{-17}	$8.4.10^{-1}$
2.10^4	293	2.10^{-4}	$8.4.10^{-1}$
1.10^4	291	7.10^{-3}	$8.4.10^{-1}$
5.10^3	290	$4.5.10^{-2}$	$8.35.10^{-1}$
2.10^3	283	$1.7.10^{-1}$	$8.1.10^{-1}$
1.10^3	273	$3.3.10^{-1}$	$7.65.10^{-1}$

It has already been pointed out that in the case of partial or complete insolubility the polymer is not usually present as a compact phase but in a swollen gel state. On similar groups as above, there will be, below T_{cr} an equilibrium between the solvent

molecules in the polymer and those outside. The composition of the concentrated phase however is not very much influenced by changes of the molecular weight, as was shown above. The structure of such a gel is that represented by Fig. 1 d, p. 157. It is very probable that the coherence in this type of gel is much less than in that represented by Fig. 1c; perhaps it will have the character of a highly viscous fluid to a greater extent.

It can now be understood that for many high-polymers the so-called "Bodenkörperregel" [1], by which the concentration of the diluted phase is dependent on the relative quantity of the solid component in a saturated solution, is valid [2]. Referring to Table 1 it is clear that more is dissolved on adding new quantities of polymer to the solution because then more of the smaller (more soluble) fractions become available. Furthermore it may happen that the solubility of long molecules is increased by the smaller ones; thus, the concentration of palmitic acid dissolved in CCl_4 can be doubled, when 1% of lauric acid is added.

In contrast to this, the "Bodenkörperregel" does not hold for polymers with molecules of uniform size (certain natural polymers, see p. 23) so that in this way it is even possible to control the uniformity of a product [3].

It will be clear from the above considerations, that the relation between the molecular size and the solubility can be followed by means of the precipitation method which is described in full on p. 144, being one of the ways of determining the molecular weight of polymers. In the experiments [4] illustrated by Fig. 5 the dissolved polymers were titrated with a nonsolvent; on the ordinate, γ indicates the volume percentage of the latter, necessary to cause precipitation.
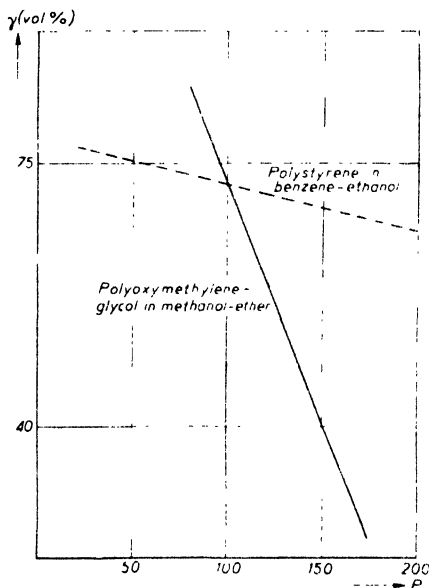
Fig. 5. The precipitability of dissolved macromolecules increases with the polymerisation degree. (Polystyrene: scale P/50). (4.4% solution).

d. The influence of the chemical constitution

In all problems of dissolution the mutual interaction between the various atomic groups of the polymer and the solvent plays a part.

---

[1] To be translated as: "The rule about the mass, lying on the bottom". Cf. for this rule, OSTWALD, *Kolloid-Z.*, 41 (1927) 163; 43 (1927) 249. A. VON BUZAGH, *ibid.*, 41 (1927) 165, 169; 43 (1927); 215; 48 (1929) 33; 50 (1930) 65.
[2] In contrast to the GAY-LUSSAC rule for low-polymers, where the solubility is independent of the quantities of solvent and dispersed substance.
[3] J. H. NORTHROP, *Am. Rev. Biochem.*, 7 (1938) 37.
[4] E. L. LOVELL and H. HIBBERT, *J. Am. Chem. Soc.*, 61 (1939) 1916. The data for polystyrene were obtained by H. STAUDINGER and W. HEUER, *Z. physik. Chem.*, A 171 (1934) 129.

It is impossible at the time being to give quantitative rules for this interaction; usually the empirical rule: "like molecules to like" is adhered to..

Van Arkel[1] has endeavoured to refine this rough rule showing that, in order to obtain ideal solvability, the three parts of the cohesion energy for the solvent — the dipole energy, the induction effect and the dispersion effect — should each almost equalize those for the polymer. We possess no experimental material to check this theory but we may refer to recent work by Gee[2], who found a relation between the swelling power of different liquids and the cohesive energies involved. As the complete treatment of this subject would lead us too far from our present subject, we will only point to some data[3] about heat-development, when soaking rubbers of various polarity in the non-polar vapour of benzene, indicating that, according to expectation, an increase of the polarity of the rubbers leads to a decrease in the heat-development. At first sight it is surprising that gutta-percha evolves much less

TABLE 3

HEAT EVOLVED WHEN SOAKING RUBBERS OF VARIOUS POLARITY INTO BENZENE-VAPOUR

Kind of rubber		Dipole moment of the polar group	Heat developed cal/g rubber
plantation rubber	$\begin{array}{c} C \\ | \\ -C-C=C-C- \end{array}$	0	593
gutta-percha	*ibid.* (crystallised) stereoisomer, see p. 662)	0	329
chlorinated rubber	$\begin{array}{c} CH_3 \\ | \\ -C-C=C-C- \\ |\ \ |\ \ |\ \ | \\ Cl\ Cl\ Cl\ Cl \end{array}$	unknown	143
Neoprene	$\begin{array}{c} Cl \\ | \\ -C-C=C-C- \end{array}$	$2.2 \cdot 10^{-18}$	61
Thiokol A	$\begin{array}{c} -C-C-S-S- \\ \ \ \ \ ||\ \ || \\ \ \ \ \ S\ \ S \end{array}$	$3 \cdot 10^{-18}$	43

heat than the chemically identical (*stereo*chemically different, but still non-polar) substance rubber. It will be shown on p. 166 that this is due to crystallisation phenomena.

[1] A. E. van Arkel, *Chem. Weekblad*, 31 (1934) 490.

[2] Gee has connected the swelling power with the so-called cohesive energy density V, where E = cohesive energy and V is molecular volume. It thus expresses the energy required to separate the molecules per cm^3 to a distance sufficiently great for them to be out of each other's sphere of attraction. The function between the degree of swelling and the difference between the cohesive energy densities of solvent and high polymer shall show a maximum and this has actually been found experimentally for rubber. Although there are still deviations in the details, the theory developed opens wide perspectives for a quantitative understanding of the phenomena involved.

[3] Y. Tanaka and co-workers, *Rubber Chem. and Technol.*, 10 (1937) 708.

In accordance with Fig. 1b p. 157 a decreased swelling can be found in non-polar solvents, the more polar the rubbers are, as appears [1] from the three righthand examples of Fig. 6. On the contrary three samples on the left hand vary only in their polarisability, as they contain an increasing number of benzene groups per ground-molecule. This evidently also leads to a higher resistance against swelling; the macromolecules seem to attract each other more strongly as a consequence of this mutual polarisation.

It is interesting to observe in Fig. 6 that, although the strongly polar ethyl alcohol as a whole is a bad swelling medium, its solvent power on the more polar rubbers is definitely better than that on the less polar ones. The scale of Fig. 6 is too small to demonstrate this clearly and there-
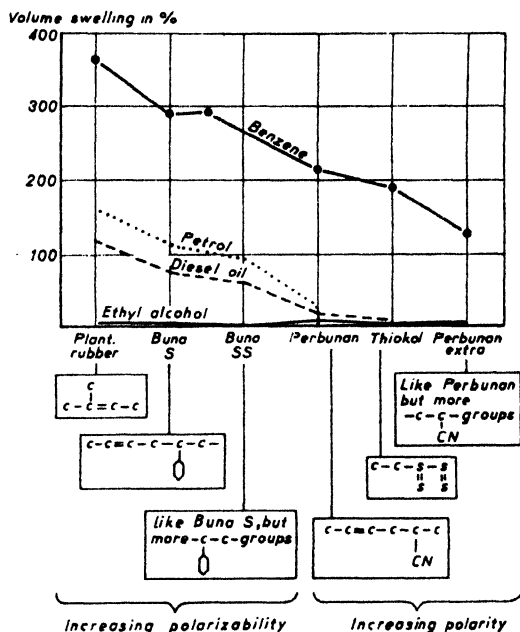
Fig. 6.   Swelling of rubbers with increasing polarity in various solvents (8 weeks at 20° C).

fore the following data may be mentioned:

	% swelling in ethyl alcohol
plant rubber . .	2
Buna S	1
Buna SS . . .	1
Perbunan . . .	10
Perduren . . .	10

Thus, one is nearly always in a position to find, for any polymer, a suitable solvent by an adequate choice of the polarity relations. This may be demonstrated [2] by Fig. 7, where the osmotic pressure relations for ethylcellulose in various solvents are shown. It appears that for the three

Fig. 7.   Influence of solvents of different polarity on the osmotic pressure of ethylcellulose.

[1] G. EHLERS, Kunststoffe, 31 (1941) 422.
[2] E. STEURER, Kolloid-Z., 96 (1941) 333.

polar solvents benzene-ethyl alcohol, chloroform, and dioxane, p/c (p = osmotic pressure, c = concentration) has the same value at infinite dilution, corresponding to a molecular weight of 48 000. This indicates that complete solution, i.e., complete separation of the particles from each other, then takes place. On increasing the concentration, p/c increases for the various solvents to a different extent, showing their varying solving power according to their individual polarity relations.

The behaviour of the non-polar solvents toluene and benzene is completely different. The low p/c values indicate a clogging together of the particles; for toluene at zero concentration a molecular weight of 141 000 can be calculated, indicating the formation of groups of three molecules.

In a similar way as described on p. 161, it is possible to demonstrate the influence of the polarity of a liquid on its solving power by titration. Thus ERBRING[1] showed that, according to Fig. 8a, the higher the number of C-atoms in the alcohol (thus becoming less polar), the greater is the quantity necessary to cause opaqueness in certain (mostly non-polar) solutions of high-polymers, showing their increasing dissolving power. In Fig. 8b the reverse situation is met with; here the polarity relations evidently are such that, on decreasing the polarity of the alcohol, a smaller quantity is required to bring about opaqueness, showing their decreasing dissolving power[2].
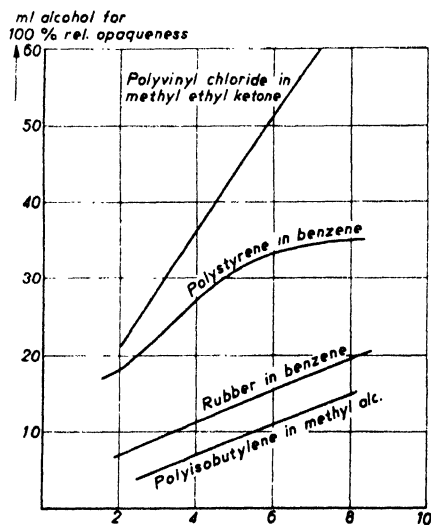
a) The less polar alcohols (higher number of C-atoms) are better solvents.
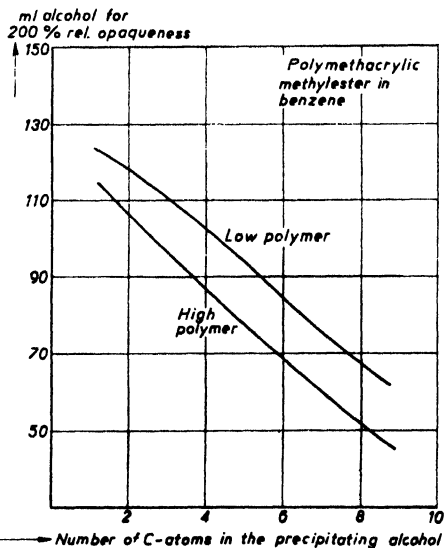
b) The less polar alcohols are poorer solvents.

Fig. 8. Precipitation experiments.

e. The influence of interlinking

The fact that the interlinking of chain-molecules is the most effective way of

[1] H. ERBRING, *Kolloid-Z.*, 90 (1940) 256.

[2] W. L. H. MOLL, *Kolloid-Z.*, 85 (1938) 335 has shown that a relation exists between the electric constant of the solvents and their surface tension.

making insoluble but swelling materials has already
been discussed. The earliest example investigated
quantitatively is that of polystyrene, interlinked by
means of p-divinylbenzene. The most important con-
clusion[1] to be drawn from these experiments is that
1 molecule of the latter in 40 000 molecules of styrene
is sufficient to make the polystyrene insoluble. GEE
has demonstrated the relation illustrated in Fig. 9,
between the volume of liquid Qm absorbed per unit
volume of rubber and the mean molecular weight M_c
between the junction points of the network.

Again the difference between good and bad sol-
vents appears.

It was pointed out on p. 23 that, in a certain
respect, there is a continuous transition from inter-
linked chains to three-dimensionally polymerised mole-
cules, the number of bridges steadily increasing. A
marked difference remains, however, in the solubility
and swelling properties. Substances with interlinked
chain-molecules become completely insoluble with a
very small degree of interlinking, their swelling power
then how-
ever being
very great.
With three-
dimensional
polymers, on the other hand, the polymeri-
sation can be carried on to a high degree,
some soluble components still remaining,
the swelling however being small.
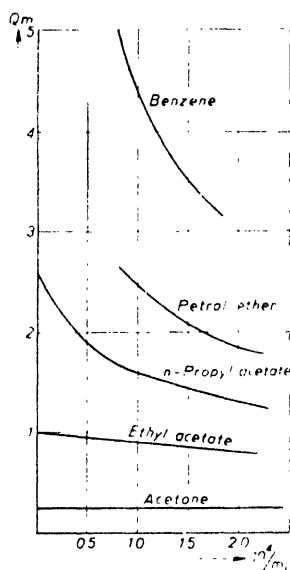
Fig. 9. Degree of swelling as a
function of the molecular weight
M_c between the junction points
in a vulcanised rubber network

f. The influence of crystallisation

When high-polymers crystallise, which
is often the case with chain-molecules, a
state of lower potential energy is obtained
compared with the amorphous form. In
the equation (2) on p. 154 U_{PP} will then
be increased making it more difficult to
dissolve the substance.

The effect has been mentioned al-
ready on p. 160, where amorphous cellulose
was shown to swell in water, whereas cry-
stalline cellulose is insoluble. It can also
be demonstrated by comparing natural rub-
ber with gutta-percha, these substances

Fig. 10. Increase of solubility of polyethylene
on heating.

[1] H. STAUDINGER, Trans. Faraday Soc., 32 (1936) 323.

being stereoisomers (cf. p. 662). The heat of crystallisation of gutta is 8 cal/g, compared with only 4 cal/g for rubber. For such solvents, where $\triangle U$ from eq. (1) is of the same magnitude as $T \triangle S$, this may just be decisive as to whether or not dissolution will take place. Thus, one encounters the fact that rubber and benzene form a solution, whereas gutta can only be dissolved after heating above its melting point. In a similar way the degree of swelling of these two polymers is influenced by crystallisation.

Another example [1] in which the influence of crystallisation on the solubility appears, is given in Fig. 10, relating to polyethylene. On increasing the temperature, a noteworthy increase of the solubility occurs at 70°C in a series of solvents. In this connection it is interesting to remark that at 70°C incipient melting of some of the crystals occurs, which is complete at 120°C.

§ 3. VISCOSITY IN DILUTE SOLUTIONS

a. Theoretical basis

The interpretation of viscosity measurements [2] is usually founded on the EINSTEIN formula, already discussed in Vol. I by KRUYT:

$$\eta = \eta_o (1 + K\varphi) \qquad (6)$$

η = Viscosity of the solution
η_o =   ,,   ,,   ,,   solvent
φ = Volume fraction of the dispersed substance
k = constant

This formula was derived from considerations of the energy dissipation in very dilute solutions, containing *rigid globular* particles, being *large* with respect to the solvent molecules. Under these ideal conditions EINSTEIN calculated k to be 0.025. The proportional increase of η with the volume concentration, is usually expressed as follows:

$$\triangle \eta_{sp} = \frac{\eta - \eta_o}{\eta_o} = \frac{\eta}{\eta_o} - 1 = \eta_r - 1 = K\varphi \qquad (7)$$

where $\triangle \eta_{sp}$ is denoted as *specific viscosity increase*
and η_r ,,   ,,   as *relative viscosity*.

It is important to repeat here that the diameter of the globules does not appear in the formulae, indicating the (experimentally verified) fact, that η_{sp} is independent of the particle size (for rigid globules).

We will base our considerations on the equation (8):

$$\triangle \eta_{sp} = 0.025 \, V_o \, C_v \qquad (8)$$

where V_o = rheological "voluminosity"
C_v = volume percentage of the intake of the dispersed substance,
using the *rheological* [3] "*voluminosity*" V_o — for reasons of convenience to be called

[1] E. L. MIDWINTER, *Inst. Plast. Ind. Trans.*, Jan. (1947) 21.
[2] Cf., for general surveys: E. C. BINGHAM, *Plasticity and Fluidity*, New York, 1922; E. HATCHEK, *The Viscosity of Liquids*, London, 1928; R. HOUWINK, *Elasticity, Plasticity and Structure of Matter*, Cambridge, 1937; First and Second Report on Viscosity, *Roy. Soc. Sci.* Amsterdam, 1938 and 1939; H. MARK, *The General Chemistry of High-polymeric Substances*, Amsterdam, 1940; W. PHILIPPOFF, *Die Viskosität der Kolloide*, Dresden, 1942.
[3] Rheology is the science of flow (Greek, $\varrho \varepsilon \iota$ = to flow).

voluminosity — as a measure, indicating the extent to which the particles deviate [1] from the ideal EINSTEIN conditions. In all cases, where V_o differs from unity the particles may be either too small, non-globular, non-rigid or swollen coils, the latter causing their volume in the solution to correspond no longer with the original intake.

STAUDINGER has introduced the conception of the *limiting-concentration* [2], by which is meant the concentration, above which η does no longer increase proportionally to c_v. Strictly speaking, this is incorrect [3], but this expression is still in use for practical applications.

This viscosity-concentration relation is shown in Fig. 11 for a series of nitrocellulose of various polymerisation degrees. For all members of a certain polymer the limiting concentration is found at the same specific viscosity. For various polymers this characteristic viscosity differs slightly, lying however in the neighbourhood of 0.1.

In accordance with STAUDINGER *we will consider solutions below the limiting concentration as dilute solutions.* Some years ago it was not possible to derive conclusions

Fig. 11. Viscosity concentration relation for nitrocellulose of different polymerisation degrees

from viscosity-data above STAUDINGER's limiting-concentration. Now this situation has changed and therefore a special section will be devoted to these so-called concentrated solutions.

It has already been remarked above [3] that η does not increase absolutely proportionally to C_v. Since η is often also dependent on the shearing stress applied (see p. 168, Fig. 12) it may, strictly speaking, only be considered at zero concentration and at zero shearing stress. PHILIPPOFF [4] has introduced on this basis a universal constant for the viscosity, namely the *viscosity-constant*, defined as:

---

[1] Strictly speaking this procedure is incorrect and unscientific, because V_0 is defined as the voluminosity at lim. $c_v \rightarrow 0$. It is however correct to study the changes of "a constant V_0" by increasing the concentration; a second question is the way in which this shall be interpreted. We will follow the investigators in this field by expressing by means of Vo the way in which the volume of the particles *seems* to be altered when changing the concentration; accordingly V_0 will be denoted as the rheological voluminosity.

[2] German: Grenzkonzentration. Cf. H. STAUDINGER, *Die hochmolekularen organischen Verb.*, Berlin, 1932.

[3] There is no straight part at the beginning of the $\eta - c_v$ curve and its slope can only be indicated by the tangent at the point $c_v = 0$. In practical experiments, however, the beginning of the curves appears to be sufficiently flat to consider the curves as straight lines, and to determine the limiting concentration with a reasonable accuracy.

[4] W. PHILIPPOFF, *Kolloid-Z.*, 98 (1942) 90.

$$[\eta] \quad \left(\frac{d\,\eta_r}{d\,C_v}\right) \; Cv \longrightarrow 0$$

$D \longrightarrow 0$ wherein D — velocity gradient[1].

In America $[\eta]$ has been introduced under the name of *intrinsic viscosity* by KRAEMER[2] for the value at infinite dilution, based on g per 100 cm[3] of solution. It can be shown[3] to be equal to

$$[\eta] = \lim_{c \to 0} \frac{\eta_{sp}}{c} = \lim_{c \to 0} \frac{\ln \eta_r}{c} \tag{9}$$

In most cases of quasi-viscosity[4] there is an upper limit of τ, above which η becomes independent of it again. This is shown, in the flow diagram of Fig. 12, borrowed from PHILIPPOFF[5] where D and τ are plotted on a logarithmic scale. Thus, at very small stresses one can speak of a viscosity coefficient (η_o), but at very high stresses one finds another coefficient (η_∞). Between these two the region of quasi-viscosity or structural viscosity is encountered. The appearance of a constant value for η_∞ may be due to the fact, that above a certain stress long-shaped particles are completely oriented so that no further structural changes take place on increasing τ still more.
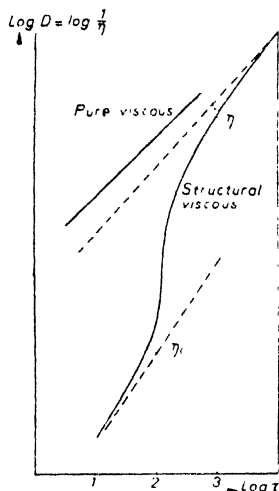
Fig. 12. Flow diagram by PHILIPPOFF.

b. Applications

b. 1. *Molecular weight determinations*

One of the most important applications of the viscosity measurements of high-polymers is that we are in a position to determine, in this way, the molecular weights of certain substances. A complete treatment of this subject is given on p. 140. Reference will only be made here to the fact that STAUDINGER[6], the pioneer in this field, uses the following formula for chain-polymers:

$$\eta_{sp} = KCM \tag{10}$$

K = a constant, characteristic for each
      series of polymer homologues
M = molecular weight; C = concentration.

There is, however, evidence that this formula represents only a special case (for $n = 1$) of the more general KUHN formula (11) to be discussed now.

---

[1] D is defined as $\dfrac{dv}{dy} \cdot \dfrac{1}{\eta}$ for the case of laminar flow, see KRUYT in Vol. I of this book.

[2] E. O. KRAEMER, *Ind. Eng. Chem.*, 30 (1938) 1200. See also this book p. 140.

[3] R. H. EWART, *Advances in Colloid Science*, Vol. 2, New York, 1946, 216. See, on this subject, L. H. CRAGG, *J. Colloid Sci.*, 1 (1946) 261.

[4] One speaks of quasi-Viscosity, when η changes with increasing shearing-stress τ.

[5] W. PHILIPPOFF, *Die Viskosität der Kolloide*, Dresden, 1942.

[6] H. STAUDINGER, *Trans. Faraday Soc.*, 32 I (1936) 108.

b. 2. Determination of the form of the molecules

Restricting ourselves chiefly to viscosity measurements, for the purpose of studying the form and further particulars of the molecules, we will start from the formula

$$[\eta] = KCM^n \tag{11}$$

where n is a constant,

derived on theoretical grounds for dispersed coils in general. From calculations by KUHN [1], n may vary between 0.5 and 0.9 for the ideal statistical coil; HUGGINS [2] has suggested that $n = 1$ for such coils. For stiff rod-like structures, HUGGINS finds $n = 2$.

A further refinement leads to two different mechanisms of flow:

a. The polymer molecule keeps the solvent entrapped (immobilised), forming more or less a sphere, to which the EINSTEIN equation (6) can be applied.

b. The liquid is free and flows between the polymer molecule as through the holes of a sieve (so-called washed through mechanism).

For mechanism a., HULBURT and co-workers [3] find $n = 0.5$, while HUGGINS [4] calculates $n = 0.9$ to 1 for mechanism b.

The experimental material available (Table 4) shows that n varies between 0.6 and 1.1, giving no definite distinction between the mechanisms of flow mentioned, although the right order of magnitude is found. It seems reasonable to suppose that the compactness of the coils depends on the mutual attraction between the groups on the molecule, the kind of solvent, the degree of branching, etc. Lower values of n may correspond to a more compact coil, immobilising less solvent. Further conclusions will be drawn from measurements in concentrated solutions (p. 176).

TABLE 4

VALUES FOR n IN FORMULA (11)

Polymer	solvent	n
Polyesters [5]		0.6
Methyl-methacrylate [6]	chloroform	0.85
Polyisobutylene [7]	cyclohexane	0.64
Polystyrene [8], polymerised at 60° C	toluene	0.70
,, ,, 120° C	,,	0.80
,, ,, 180° C	,,	1.10
Cellulose acetate [9]	acetone	0.67
Rubber [10]	toluene	0.67

[1] W. KUHN, *Kolloid-Z.*, 68 (1934) 2.
[2] M. L. HUGGINS, *J. Phys. Chem.*, 42 (1938) 910; 43 (1939) 439.
[3] H. M. HULBURT et al. *Ann. New York Acad.*, 44 (1943) 371.
[4] M. L. HUGGINS, *J. Phys. Chem.*, 42 (1938) 910; 43 (1939) 439.
[5] R. HOUWINK, *J. prakt. Chem.*, 157 (1940) 15.
[6] G. V. SCHULZ and A. DINLINGER, *J. prakt. Chem.*, 158 (1944) 136.
[7] P. J. FLORY, *J. Am. Chem. Soc.*, 65 (1943) 372.
[8] T. ALFREY, A. BARTOVICS and H. MARK, *J. Am. Chem. Soc.*, 65 (1943) 2319.
[9] H. BARTOVICS and H. MARK, *J. Am. Chem. Soc.*, 65 (1943) 1901.
[10] W. C. CARTER, R. L. SCOTT and M. MAGAT, *J. Am. Chem. Soc.*, 68 (1946) 1480.

It will be clear from these considerations that in many cases rather large quantities of solvent are assumed to be present in the swollen particles. This does not mean, however, that the older conceptions of thick solvent mantles around the polymer molecules should be maintained. Presumably the picture[1] shown in Fig. 13 gives the right representation for the possibilities involved, discriminating between *true* solvation, intramicellar and intermicellar swelling based on a conception of EIRICH and MARK[2]. It must however be stressed that in the two latter cases only a *mechanical* immobilisation (enclosure) of solvent is considered, the true solvation including only layers of a few molecules thickness[3]. It is unnecessary to say that the formation of clusters, which may contain some thousands of molecules[4], will strongly increase the opportunity of immobilising liquid.

| linear   globular | linear   globular | linear   globular |
| True solvation | Intramicellar swelling | Intermicellar swelling |
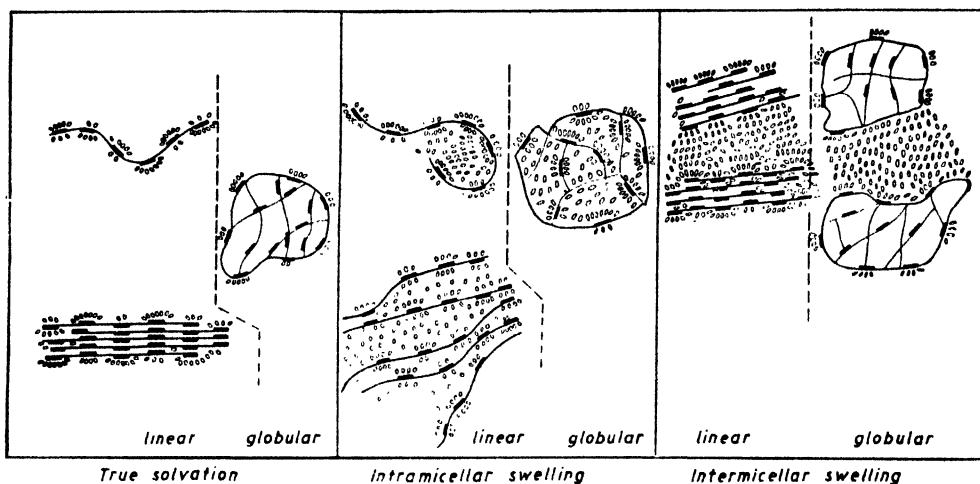
Fig. 13. Possible causes for $Vo > 1$

If there were rigid rods instead of coils — STAUDINGER's earlier supposition — n should be 2 according to KUHN and HUGGINS. The above mentioned results do not therefore confirm this frequently encountered opinion about the rigid rods, it being as a rule[5] improbable on also other grounds.

[1] R. HOUWINK and K. H. KLAASSENS, *Kolloid-Z.*, 79 (1937) 138.
[2] F. EIRICH and H. MARK, *Ergebn. exakt. Naturwiss.*, 15, Berlin, 1936.
[3] It will be necessary in future to have a criterion by which molecules shall be reckoned to be still solvated. It seems logical to consider that the solvation layer extends to points at whici. the binding energy of the molecules is of order of magnitude of the kinetic energy of the solvent molecules.
[4] H. A. STUART, *Z. Physik*, 63 (1930) 533.
[5] Perhaps for certain polymers where the possibility of rotation between the chain elements is practically lacking, this structure may be valid. So it may be that polystyrene belongs to this type, due to the mutual steric hindrances of its very voluminous C_6H_5-side-groups. Cf. J. H. DE BOER, *Trans. Faraday Soc.*, 32 (1936) 10.

TABLE 5

VALUES FOR THE VOLUMINOSITY Vo IN VARIOUS SOLS [1]

Sol	particulars about the dispersed particles	Vo		dispersion medium
		globular particles	long-shaped molecules	
Paraffin oil emulsion . .	2 to 220 μ	1.0	—	water
Rubber latex	1 to 3 μ	1.02	—	water
Cresol-formaldehyde resin	low polymeris. degree	1.0	—	acetone
Phenol-formaldehyde resin	,,	1.3	—	acetone
Sulphur	—	1.8	—	water
Polystyrene	mol. wt 2400	—	2.7	benzene
,, ...	,, 7500	—	5.6	benzene
,, ...	,, 23000	—	15.4	tetralin
,, ...	,, 120000	—	80.5	tetralin
,, ...	,, 280000	—	205.0	tetralin
Acetylcellulose	,, 1810	—	3.4	m-cresol
,, ...	,, 6400	—	13.2	,,
,, ...	,, 13600	—	31.3	,,
Viscose	28 days matured	—	53.3	7% NaOH
,, ...	21 ,, ,,	—	61.5	7% NaOH
,, ...	14 ,, ,,	—	74.0	7% NaOH
,, ...	7 ,, ,,	—	98.5	7% NaOH
,, ...	1 ,, ,,	—	175.0	7% NaOH
Rubber	lightly masticated	—	90.6	chlorinated benzene
	heavily masticated	—	15.0	chlorinated benzene
Cotton	nitrated	—	900	butylacetate

Considering now the voluminosity Vo from (8), it appears from Table 5 that, in practice this constant varies between 1 and about 1000.

Paraffin emulsions and rubber latex, the particles of which can be checked microscopically to be globular, actually show values in the neighbourhood of 1 so that these seem to be sufficiently rigid to fulfill the elementary EINSTEIN conditions. Sulphur sol has a slightly higher value of 1.8; perhaps here the influence of the ionic atmosphere plays a part (see Chapter VII, OVERBEEK and BUNGENBERG DE JONG).

The cresol- and phenol-formaldehyde resins also show Vo-values in the neighbourhood of 1, in accordance with their globular shape. The fact that the values found are not exactly 1 may be due to their being somewhat swollen in the solvent.

The polymers with chain-molecules show higher values, depending on their polymerisation degree. Although, as already remarked, it is possible that in certain cases extended chain-molecules as such may be responsible for this high viscosity, the modern view is that they form coils or in certain cases swollen net-fragments, the *magnitude* and *compactness* of which decide the values of Vo.

[1] Data taken from various experiments scattered through literature and summarised by H. L. BREDEE and J. DE BOOYS, Kolloid-Z., 79 (1937) 31, 43; 91 (1940) 39; 99 (1942) 171; R. HOUWINK and K. H. KLAASSENS, Kolloid-Z., 79 (1937) 137; 99 (1942) 160.

That their *magnitude* has an influence, follows from the data for polystyrene and acetylcellulose, indicating that η increases with the molecular weight. At first sight this cannot be understood easily, coils being expected to be more or less globular and therefore η_{sp} should be independent of their diameter (which is actually found for the paraffin emulsion, in Table 5). The solution of this difficulty, however, was given by KUHN[1], who showed that the coils have a bean-like form and who calculated that η_{sp} increases with $M^{0.5 \text{ to } 0.9}$

The decrease of Vo, when maturing viscose or when masticating rubber show, that for both these products the technical operations involved are accompanied by a breakdown of the macromolecules.

Mean velocity of flow
$V_\bullet \dfrac{4\varphi}{\pi R^3}$ in sec⁻¹

Mol wt 86,000
174,000
358,000
Mol. wt 1,025,000

Shearing stress in dynes/cm²
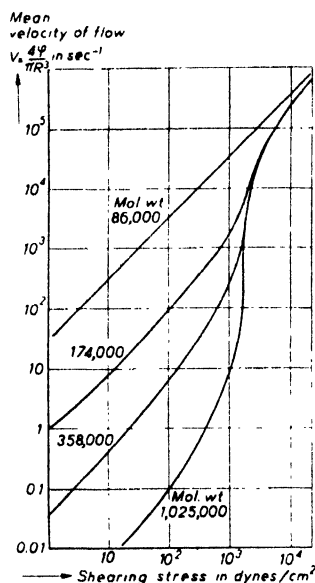
Fig. 14. Flow curves for nitro-cellulose of various polymerisation degrees (obtained by bleaching).

In many cases useful additional information about the particle shape can be obtained from the flow curves (cf., p. 168). In Fig. 14 a series of these[2] for nitrocellulose of various polymerisation degrees is shown, and it appears that the higher the molecular weight, the more the curve deviates from a straight line. The cause of this is probably that long rigid particles are more easily oriented in the field of flow, leading to a more pronounced structural viscosity. It is however dangerous to draw definite conclusions from these curves, because other causes also may lead to structural viscosity, e.g., the mutual loosening of micelles (p. 174).

Returning to the low Vo values, for the paraffin emulsions and for rubber latex of about 1, this is all the more remarkable because *in the globular particles of these emulsions the chain-molecules are still present*, being, however, packed closely together in a thin envelope[3]. *Emulsifying a polymer therefore is the most simple way of preparing dispersions of long chain-molecules with a low viscosity*, and on p. 44, 45 we have encountered this method, applied in order to enable stirring during polymerisation. The extent to which the situation differs when the chain-molecules are liberated from their envelope by dispersion in an organic solvent appears immediately from the Vo value for lightly masticated rubber of $90 \cdot 6$. Thus, one is in a position to change the viscosity of such types of dispersions between wide limits by varying only the particle shape.

It will now be clear that viscosity measurements are often an adequate means of deciding whether the molecules are round or not. When, however, Vo is low, doubt can arise about the interpretation, because chain-molecules might be present with a polymerisation degree which is insufficient for a high viscosity to be obtained.

[1] W. KUHN, *Kolloid-Z.*, 68 (1934) 2. Very probable dimensions of this bent ellipsoid are axes, related to each other as 6 : 2. 3 : 1. The density of this ellipsoid decreases on moving from the centre, leading to a greater relative volume for larger coils and thus to the viscosity increase mentioned. For the case of rigid rods KUHN calculates that η_{sp} increases with M^2.

[2] W. PHILIPPOFF and K. HESS, *Z. phys. Chem.*, B 31 (1936) 237.

[3] The globules are surrounded by a film of emulgator; in the case of rubber latex this is composed of proteins.

This is demonstrated by the example of polystyrene with a mol.wt. of 2 400, which has a *Vo* of only 2·7. In such a case a decision can often be made by carrying the polymerisation to a further stage. When the particles are globular, practically no change of *Vo* is then to be expected, but when chain-molecules are present, *Vo* must increase. The results of such an experiment are illustrated in Table 6, where the *Vo*-increase on polymerisation of globular phenol-formaldehyde resin is compared with that of polystyrene.

TABLE 6

Vo-INCREASE ON POLYMERISATION FOR RESINS WITH GLOBULAR AND WITH CHAIN-MOLECULES RESPECTIVELY

Phenol-formaldehyde resin (globular); 2% in acetone [1]		Polystyrene (chain-molecules); 1% in tetralin [2]	
Reaction time	Vo	Mol. wt	Vo
20 min at 100° C	1.1	312	1
60 ,,	1.3	7600	6
90 ,,	1.7	120000	88
		600000	440

b. 3. The chemical structure

Considering the influence of the chemical structure of the highpolymer particles and of the solvent molecules on solubility (p. 164), there is little wonder that these effects will also have a bearing on *Vo*. In Fig. 15 the change of the viscosity of a series of acetylcelluloses, possessing an increasing number of acetyl groups, is illustrated [3] for two different solvents. This influence of the chemical constitution
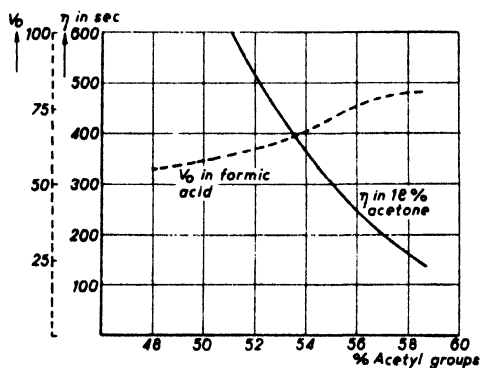
Fig. 15. Influence of the chemical constitution of a cellulose derivative on the viscosity in solution.
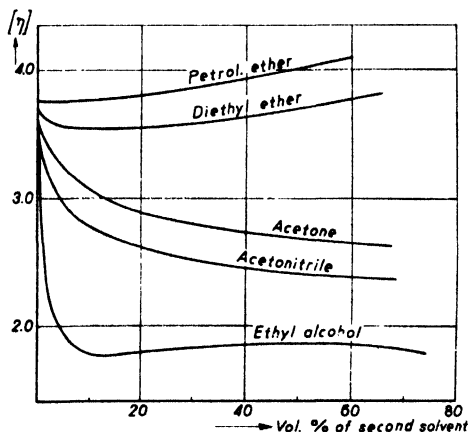
Fig. 16. Viscosity of ethylcellulose in benzene (0.75%) when adding a second solvent (petrol-ether etc.).

[1] K. H. KLAASSENS and R. HOUWINK, *Kolloid-Z.*, 76 (1936) 217.
[2] H. STAUDINGER and W. HEUER, *Die hochmol. organ. Verb.*, Berlin, 1932.
[3] S. ROGOWIN and M. JOFFE, *J. allgem. Chem.*, VII 69 (1937) 2167.

may be connected with the different compactness, which the particles obtain in the solvents. It must be said, however, that the tendency to form micelles may also play a part here. This was specially stressed by SAUTER[1], who gave the following data on the viscosity changes of a 0.75% ethylcellulose solution in benzene, on adding a second solvent (petrol ether, etc.), keeping the concentration of the ethylcellulose constant. Here the non-polar solvents (petrol ether, diethyl ether) cause a viscosity increase which is ascribed, by him, to the formation of micelles whereas the opposite effect of the three polar solvents is ascribed to their dispersing tendency with regard to the micelles.
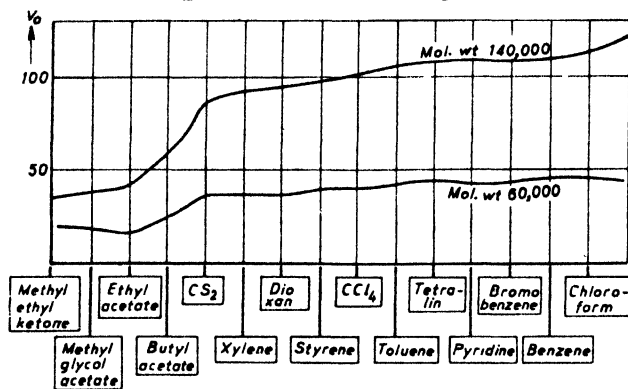
Fig. 17. V_o -values for polystyrene in different solvents.

To give another example, the specific influence of some solvents is demonstrated[2] for two kinds of polystyrene of different molecular weight (Fig. 17). The trend is the same for both materials, although no definite quantitative relation between the polarities of solvent and dissolved particles can be traced.

b. 4. The presence of loose micelles

Viscosity measurements have also been applied by STAUDINGER to decide[3] whether the macromolecules are present as *loosely* cohering *micelles* in a solution. In this case their coherence is gradually broken by heating, which means that the network, originally formed by secondary bonds, is demolished, leading to a sharp viscosity decrease[4]. On cooling however the net is restored again, leading in extreme cases to the formation of a gel. Thus, it is found for a 7% solution of oleic acid in KOH that $\frac{\eta\ 20°C}{\eta\ 60°C} = 14$, indicating the presence of micelles, whereas for polystyrene solutions, with their individual chain-molecules, this value is only 1.3. There are also however, factors which cause a viscosity increase on heating, as may be taken from the example[5] of polyacrylic acid dissolved in water, where the ratio $\frac{\eta\ 20°C}{\eta\ 60°C}$ is found to be about 0.97. Such a behaviour may be explained by assuming that certain

[1] E. SAUTER, Z. phys. Chem., A 190 (1942) 16.

[2] H. STAUDINGER and W. HEUER, Z. phys. Chem., A.171 (1934) 129.

[3] H. STAUDINGER, Die Hochmol. organ. Verb., Berlin, 1932.

[4] The solvent itself and other simple liquids also, all show a viscosity decrease on heating due to the increasing separation of the molecules from each other. This effect however is very small;

$\frac{\eta\ 20°C}{\eta\ 60°C}$ being about 1·04.

[5] H. STAUDINGER, Die Hochmolekularen organischen Verbindungen, Berlin, 1932, p. 347.

rotations in the molecule are still restricted at the lower temperature but become free at the higher temperature and that the extended form has a lower potential energy.

The experimental observation may therefore be the net result of increasing and decreasing factors, often making conclusions difficult to draw from such measurements.

Another way of obtaining indications about the presence of loosely cohering micelles is by investigating the flow curve on applying various shearing stresses. Curves of the type shown on p. 172 may be the consequence of a loosening of the bonds between the particles by the stresses applied. As has already been remarked on p. 172, however, the exact interpretation is still very complicated.

b. 5. The influence of polydispersity

As it was shown on p. 44 that high-polymers usually contain molecules of different size, the molecular weight always representing therefore a mean value, it is important to decide the extent to which the viscosity of solutions is dependent on this polydispersity.

In this respect it must first be stated that, for mixtures of chemically identical substances, the viscosity is found to be [1] composed additively so that one can write:

$$\eta_m = \int_{i=1}^{i=\infty} \eta_i \, x_i \tag{12}$$

where $\eta_m = \eta$ of the mixture
i = number of the component involved
x = weight fraction of a component.

For rubber, dissolved in benzene [1] the value of $[\eta]$ has been found to be in accordance with that calculated from the various components on a basis of this formula. The results are shown in Table 7.

TABLE 7

INTRINSIC VISCOSITIES OF NATURAL RUBBER FRACTIONS AND OF WHOLE RUBBER

Sample No.	Intrinsic viscosity $[\eta]_i$	Fraction of whole rubber x_i	Contribution to $[\eta]_m$
7	0.69	0.0946	0.065
6	0.90	0.0474	0.042
5	0.99	0.0321	0.031
4	1.95	0.1991	0.238
3	2.38	0.1508	0.359
2	2.60	0.1712	0.445
1	4.39	0.3048	1.338
			$[\eta]_m = 2.518$
		whole rubber	$[\eta] = 2.519$

b. 6. Thixotropy and Rheopexy

These phenomena of viscosity change under the influence of mechanical forces (KRUYT, Vol. I) are not observed to an important extent for dilute sols without

---

[1] R. H. EWART, *Advances in Colloid Science*, Vol. 2, New York 1946, 216. See, on this subject, L. H. CRAGG, *J. Colloid Sci.*, 1 (1946) 261.

electrolytes, and therefore this subject will not be discussed further. It can only be said that it makes it the more acceptable that potential curves with an energy barrier are not encountered here.

§ 4. VISCOSITY IN CONCENTRATED SOLUTIONS

It has already been pointed out on p. 167 that a few years ago it was not possible to draw important conclusions from viscosity measurements in concentrated solutions i.e., above the limiting concentration. This situation has changed since more systematic attempts have been made to cover the vast amount of experimental data by empirical formulae [1]. Only those formulae will be discussed which bring us new insights into the colloidal state of the dispersed particles.

The chief results, which will appear are:

a. that for those particles, where a loose structure is suggested (coils, loose net fragments), the voluminosity Vo diminishes with increasing concentration. It gives the impression [2] that the coils are compressed [3], but a simple explanation is that less and less solvent becomes available per coil, making it impossible for the coils to immobilise their maximum amount of solvent. The longer the chain-molecules, the more pronounced is this apparent compression; in cases where rigid globular particles are assumed, Vo remains constant.

b. that this apparent compressibility is not only dependent on the chemical structure of the polymer, but also on that of the solvent molecules.

---

[1] H. L. BREDEE and J. DE BOOYS,. *Kolloid-Z.*, 79 (1937) 31, 43; 91 (1940) 39; 99 (1942) 171; R. HOUWINK and K. H. KLAASSENS, *Kolloid-Z.*, 76 (1936) 217; 79 (1937) 138; 99 (1942) 160; H. EILERS, *Kolloid-Z.*, 97 (1941) 313; 102 (1943) 154; H. DE BRUYN, *Rec. trav. chim.*, 61 (1942) 863. In these studies all former formulae, applied to cover the viscosity—concentration relations, are checked and their merits are discussed in full.

[2] Discussing this point, attention must be drawn to the complicated character of the factor Vo. From the view point of the original calculations of EINSTEIN for globules. Vo has a bearing on the volume of the swollen particles in the solution. If their form is not globular however, this will lead to an *extra energy dissipation* in the field of flow, as has been calculated on various occasions[*]. This influence can be accounted for, according to HOUWINK and KLAASSENS[1], by writing:

$$Vo = Di \ Vr.$$
$$\text{where } Di = \text{dissipation factor}$$
$$Vr = \text{real voluminosity.}$$

This method of splitting Vo has the advantage of allowing conclusions to be drawn, from the experiments, concerning the *actual* volume Vr of the particles, no complications being produced by energy dissipation influences. By doing this, it is possible to prove, that *Di can increase but never decrease as a function of Cv*. In cases where the product $DiVr$ decreases, therefore, no doubt exists that Vr diminishes to such an extent that at least the increase of Di is overcompensated.

[3] There are still other methods to explain the compression of the particles. In this respect a conception of P. H. HERMANS, J. J. HERMANS and D. VERMAAS, *Kolloid-Z.*, 105 (1943) 199; 106 (1944) 22, 95 may be pointed to, where it is assumed that the coils form points of attachment, the number of which increases with concentration. The more points of attachment there are, the more dense will be the coil, this conception being completely in accordance with the ideas developed in the text.

---

[*] R. EISENSCHITZ. *Z. phys. Chem.*. A 158 (1932) 78. for example. calculates that for ellipsoid particles η_{sp} increases proportionally to the ratio: $\dfrac{length \ axis \ l}{cross \ axis \ d}$. Cf. also J. M. BURGERS in Second Report on Viscosity. *Roy. Acad. Sci.* Amsterdam. 1938.

c. that in contrast to the situation observed in dilute solutions, chain-polymers have a relatively smaller viscosity-increasing power in concentrated solutions than have globular particles.

The formula furnishing these conclusions in the simplest way is perhaps (13), by DE BRUYN [1]

$$Vo_c = \frac{0.8}{C} \left(1 - \left| \frac{1}{\eta_r}\right.\right)d \tag{13}$$

where Vo_c = Voluminosity at a certain concentration c.
       C = concentration in g/ml (intake)
       d = density of the dispersed substance in the dry state.

In this formula the conception of the voluminosity at a certain concentration Vo_c is introduced, in contrast to Vo, which latter is variable with concentration [2].

Plotting Vo_c C_v , expressing the so-called "rheological concentration", against Vo_c straight lines are obtained, showing that Vo_c decreases very rapidly (Fig. 18). The longer the chain-molecules, the greater is this decrease.
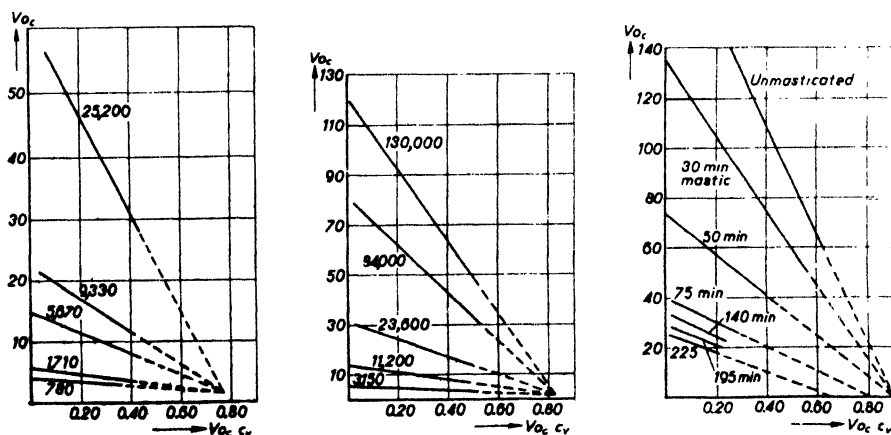
a) Polystyrene in tetralin   b) Acetylcellulose in *m*-cresol   c) Rubber in chlorobenzene.

Fig. 18. Voluminosity at a certain concentration Vo_c as a function of the rheological volume concentration for chain-polymers of different molecular weight.

The relationships appear to be rectilinear, which is understandable when realizing that Vo_c will decrease in the same ratio as the polymer is added to the solution, which means that less solvent becomes available for the chain-molecules to immobilise [3]. The more rapid Vo_c decrease in the case of longer molecules can be explained by

[1] H. DE BRUYN, *Rec. trav. chim.*, 61 (1942) 863. A formula developed by EILERS, *Kolloid-Z.*, 97 (1941) 313; 102 (1943) 154 leads to approximately the same conclusions.
[2] Vo is equal to Vo_c at zero concentration, thus being invariable by definition.
[3] Such an effect is perhaps also encountered in osmotic measurements, where a decrease of the so-called specific co-volume with concentration is found: G. V. SCHULZ, *Z. physik. Chem.*, A 158 (1932) 237.

pointing to the fact that a long molecule immobilises relatively more solvent molecules (η is proportional to some power of M, see e.g., p. 169) than a short one. Doubling the concentration, for example, reduces in both cases the number of solvent molecules available per polymer molecule to 50%; this will bring about a more important reduction of Vo_c in an absolute sense for the longer molecules.

Further, attention may be drawn to the fact that all curves converge to a point at which $Vo_c \, C_v$ is about 0.80. Theoretically this should be 0.74, corresponding to the closest arrangement which can exist on packing rigid spheres (one surrounded by 12). This may be due to the fact that the coils are not rigid spheres.

Considering now substances with rigid globular particles, such as phenol-formaldehyde resins, rubber latex and sulphur sols, it appears from Table 8 that Vo_c is independent of the concentration, being therefore equal to Vo. This can be understood because these particles do not immobilise any solvent, and are practically incompressible. Referring once more to Fig. 18, attention is drawn to the curves for

TABLE 8

RHEOLOGICAL VOLUMINOSITY AT A CERTAIN CONCENTRATION Vo_c OF GLOBULAR POLYMERS WITH RIGID PARTICLES [1]

substance	dispersion medium	concentration	rheological voluminosity Vo_c
phenol-formaldehyde resin (20 min at 100° C)	acetone	2%	1.05
phenol-formaldehyde resin (90 min at 100° C)	acetone	30%	1.02
		2%	1.06
		30%	1.6
rubber latex	water	5%	1.08
		60%	1.08
sulphur sol.	water	4%	1.8
		30%	1.8

polystyrene, mol. wt 600, and for acetylcellulose, mol. wt 3150, where constant Vo_c values are also found, although the particles are linear. Perhaps the coils are so small and thus so compact here that they behave practically as rigid particles.

Further studies have led to another important result. It was shown that *in contrast to the situation which can be observed in dilute solutions, above the limiting concentration longer chains have a relatively smaller viscosity-increasing power than shorter ones and especially is this smaller than the viscosity-increasing power of globular particles.* Summarizing this result in one illustration,
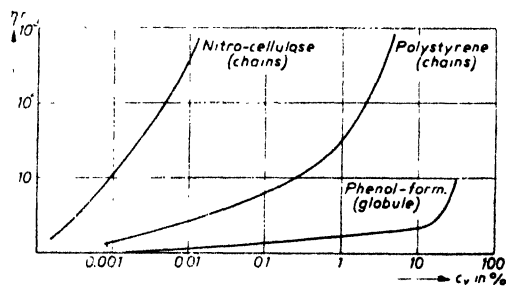
Fig. 19. The different behaviour of globular and chain-polymers with regard to their viscosity- increasing capacity in dilute and in concentrated solution.

[1] H. DE BRUYN, *Rec. trav. chim.*, 61 (1942) 863. A formula developed by EILERS, *Kolloid-Z.*, 97 (1941) 313; 102 (1943) 154 leads to approximately the same conclusions.

one obtains Fig. 19, showing clearly the relatively steeper rise in the beginning but the more gentle rise at the end for the chain-polymers.

This conclusion can be derived[1] from formula (14), also describing the viscosity-concentration relation above the limiting concentration.

$$\log \eta_r = KC_v^{a} \qquad (14)$$

$K =$ a constant with a similar meaning as Vo_c

$a =$ a constant, measuring the change of Vo_c on increasing the concentration.

In this formula the constant K measures the change of η_r, and thus of Vo_c, on increasing the concentration. It appears from Table 9 that a is 1.18 for the dispersions with rigid globular particles. For long chain-molecules η is much smaller; even values as low as 0.72 have been observed. The following analysis gives a further insight into these questions:

$$\frac{d \log \eta_r}{dC_v} = K d C_v^{a-1} = K^1 C_v^{a-1} \qquad (15)$$

Plotting $\log \eta_r$ against the volume concentration, curves of the type of Fig. 20 are found, showing that for $a < 1$ a very steep rise (high value of K) can be observed initially, which steadily decreases however.

This is in contrast to the curves for $a > 1$. Thus, one can say that for $a < 1$ the concentration increasing capacity becomes relatively smaller and smaller, which can be explained by reference to the previous discussions in which it was stated that the molecules have progressively less solvent at their disposal than they are able to immobilise. One can also express this by saying that the coils are more and more compressed. This effect is more pronounced for larger molecules. For dispersions with hard rigid globular particles however (phenol-formaldehyde resin, sulphur), the concentration increasing-power does not

Fig. 20. Curves representing $\log \eta_r$ plotted against the volume concentration

show such a tendency, which is obvious when realizing that there is no immobilised solvent present, and that there is no opportunity here for compressing any loose particles. Only in ionic dispersions, containing particles surrounded by ion clusters which develop a potential field between the particles, so that these are at a large distance from each other, can such a tendency be expected.

In concentrated sols without electrolytes the phenomenon of *thixotropy* can be found more distinctly than in dilute solutions. This is clear, since the gel state is more closely approached by concentrated sols. Thus, HÖPPLER[2] found that the velocity

[1] R. HOUWINK and K. H. KLAASSENS, *Kolloid-Z.*, 79 (1937) 138. It also appears very clearly from the important work of BREDEE and DE BOOYS (cf. note 1 on p. 176); for reasons of space this will not now be discussed further.

[2] F. HÖPPLER, *Kautschuk*, 17 (1941) 17.

TABLE 9

VALUES FOR K AND FOR α FROM FORMULA (14) FOR SOME POLYMERS

Polymer	Particulars	Voluminosity Vo from form (8), p. 166	K	α
Polystyrene	mol. wt 120000	82	15.5	0.72
in benzene	mol. wt 23000	16	5.7	0.75
	mol. wt 7500	6	4.2	0.92
	mol. wt 2400	3	3.4	1.10
Viscose in 7% NaOH	matured for 1 days	175	54	0.80
	matured for 7 days	99	34	0.80
	matured for 14 days	74	27	0.80
	matured for 28 days	53	21	0.80
Plantation rubber	not plasticised	205	56	0.77
in chlorbenzene	30 min plasticised	170	42	0.78
	75 min plasticised	49	19	0.84
Rubber-latex		1.02	1.8	1.18
Phenol-formaldehyde resin in alcohol		1.35	2.5	1.18
Sulphur sol in water		1.8	3.5	1.18

of flow of rubber sols after 2 hours standing was about 10% smaller than before, indicating that in this period an appreciable stiffening (gel formation?) took place. It may be that in this special case of rubber, the stiffening was connected with the crystallisation tendency, by which the molecules gradually drop into deeper potential energy troughs. Considering this picture, it is doubtful whether the assumption of complicated potential curves with an energy barrier are necessary in this case, to explain the rheological behaviour of electrolyte-free sols in this respect. It seems more reasonable to suppose that stirring removes certain atomic groups out of the sphere of attraction of each other, but that with these unmanageable macromolecules it takes some time to get them back into their original positions of lower potential energy.

§ 5. VISCOSITY OF ONE-COMPONENT SOLS

a. Viscosity increase on polymerisation

High polymers with such a frequency-distribution curve that large molecules are present together with small ones, can be considered as solutions of the larger molecules in a "solvent" of the smaller molecules. Both types being of the same chemical constitution, such systems were denoted in Chap. II as isosols. There are no reasons for expecting new rules for the viscosity change on increasing the concentration of the size of the large molecules, other than those encountered for heterosols.

Starting from this point of view, the viscosity increase on polymerisation can easily be understood. The transition of the smaller molecules into large ones has two consequences:

a. an increase of the *size* and
b. an increase of the *concentration* of the macromolecules.

In the case of globular molecules only [1] factor b will have a bearing on viscosity and for a description of this a formula of the type (13) or (14) may be valid.

In the case of chain-polymers however, the situation is more complicated, both

---

[1] According to EINSTEIN η is independent of the diameter of globular particles, see p. 166.

factors playing a part here. It is impossible at the present time to give a quantitive t.eatment of their combined effect.

PowELL and EYRING [1] consider that the movement of the molecules into adjacent empty equilibrium positions or *holes* is the fundamental mechanism of the flowing process. From this point of view a liquid might be conceived as a binary mixture of molecules and holes. Since the heat of activation for flow is found to be only one third of the heat of vaporisation, the latter being the energy required to make a hole the size of a molecule, the cavity must be smaller than a molecule. It is found to be about one seventh of its volume for normal liquids.

When plotting the heat of activation for flow of the paraffins against the number of C-atoms in the chain, the heat approaches a limiting value. This indicates that when the hydrocarbon becomes large, *it flows in segments of fixed size*, which are calculated to correspond to a chain length of 20-25 C-atoms. From viscosity data on molten polyesters [2] a segment length of 28-34 chain atoms was calculated; for molten polymerised sulphur, 20 atoms.

b. Temperature influence [3] on viscosity

In isosols we meet with the curious situation that neither the solvent, nor the polymer consists of molecules of uniform size. This makes a marked difference with heterosols, especially when entering the temperature region in which the associated ("solvated") smaller molecules begin to get loosened from the larger ones. Referring to p. 160, it will be clear that the smaller the molecules, the more easily will they be loosened and thus each step in rising the temperature will free an increasing fraction of molecules, changing therefore the degree of association. In this region a very pronounced viscosity decrease as a consequence of the change in the sol concentration will thus be encountered. Only when temperature has become so high that all coherence between the molecules has been loosened, may the temperature function be expected to be like that of a simple liquid. Experiments have shown that, as long as the type of association remains the same, the viscosity-temperature function can be expressed [4] by the formula (16):

$$\eta = A\,e^{-\dfrac{E}{RT_{abs}}} \tag{16}$$

where A and E are constants.

Here, especially the constant E is of importance for our purpose [5], measuring the activation energy of flow, that means the energy, necessary to destroy the mutual coherence between the molecules, thus giving information concerning the type of association in the liquid. Plotting $\log \eta$ against T, straight lines are to be expected as long as the formula (16) is valid. Fig. 21 shows however that, for most polymers, curved lines are obtained.

We will first consider the curve for glass, this having been measured over the

---

[1] See for a survey R. E. PowELL and H. EYRING, *Advances in Colloid Science* I, New York, 1942, 183.

[2] P. J. FLORY, *J. Am. Chem. Soc.*, 62 (1940) 3036.

[3] Only those isosols are considered here in which polymerisation does not proceed on heating (non-hardening type).

[4] T. ALFREY, *Mechanical Behaviour of High Polymers*, New York 1948.

[5] The constant A is related to the molecular weight and the density.

greatest region of η. We will see on p. 656 that there are arguments to consider glass as a polymer at not too high a temperature, so that a comparison with macro-molecular substances seems justified. At high temperatures it may be expected to be in the isosol state; on cooling, however, it will become an isogel.
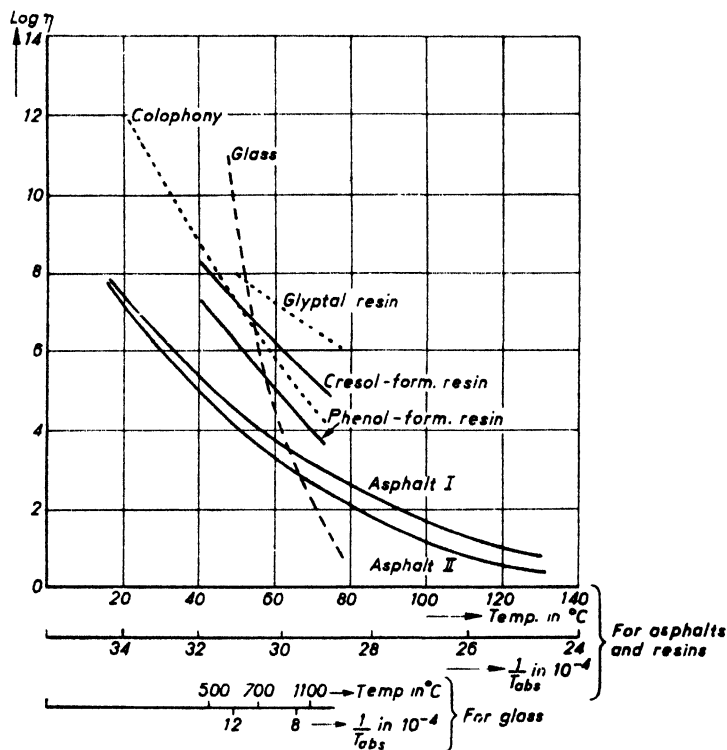
Fig. 21.   Temperature influence on viscosity for some high-polymers (note the different temperature scale for glass).

From Fig. 21 it appears that a flat hyperbola is found, the branches of which can therefore be considered, approximately, as straight lines. Applying the formula (16) to these two straight parts, different values for the constants A and E are found, as is seen in Table 10.

TABLE 10

VALUES FOR A AND E FROM THE FORMULA (16)

Polymers	log A	$E . 10^{-3}$	Substances for comparison	log A	$E \cdot 10^{-3}$
Phenol-formaldehyde resin. . .	36	64	benzene 20° C		1·06
Glyptal resin	18	38	solution of resin in amylacetate		
Copal	24	52	(diluted) 20° C.		0·2
Colophony	43	76	solution of rubber in toluene		
Asphalt	26	50	(diluted) 20° C		1·0
Glass { low temp. (500° C) . .	13	110	water (15° C)		2·2
high temp. (1200° C) . .	5	50			

At the low temperature E, has for glass, a higher value ($110 \cdot 10^{-3}$) than at the high temperature ($50 \cdot 10^{-3}$), indicating that in the first case more heat is necessary to loosen the coherence between the molecules. The transition from one straight part of the curve to the other corresponds therefore to a transition from one type of association to the other. Here the previously mentioned step-wise loosening of still more "solvent" molecules from the larger particles may be expected to take place; a gradual stopping of the Brownian movements will also take place, first of the micro- and then of the macro-motion, see Chap. XIII, p. 653. For both straight parts of the curve the coherence between the molecules is strongly influenced by temperature, as appears from the high values of K, these being of an order of magnitude completely different from those for simple liquids like benzene or water[1] (roughly 100 times higher).

For the asphalts a similar form of the curves is observed, which is in agreement with the present ideas about the structure of asphalts as highly associated polymers[2]. The curves for the resins in Fig. 21 have not been measured over a sufficient range of η to decide whether they are hyperbolic. Considering them as straight lines, B can be calculated, leading to values of the same order of magnitude as these found for glass and asphalts, showing that here too highly associated substances are present, the coherence of which is very sensitive to temperature. It is interesting to make a comparison with the values for diluted heterosols of high-polymers in a solvent. These values are roughly 100 times lower, being of the same magnitude as those for the solvent alone. This indicates, that in these heterosols no important change in the degree of association takes place in the temperature region involved. As remarked above, it is to be expected that, if the polymers themselves were heated to such a high temperature, the mutual coherence between the molecules (association) would be practically destroyed, and similar low values for B would be found.

---

[1] Even water is still slightly associated.
[2] J. Ph. Pfeiffer and P. M. van Doormaal, *J. Inst. Petrol. Techn.*, 22 (1936) 414.

VII. SOLS OF MACROMOLECULAR COLLOIDS WITH ELECTROLYTIC NATURE

J. Th. G. Overbeek and H. G. Bungenberg de Jong

Van 't Hoff Laboratory, University of Utrecht.

Department of Medical Chemistry, University of Leyden

§ 1. INTRODUCTION AND SURVEY OF THE SUBSTANCES TO BE TREATED IN THIS AND IN THE FOLLOWING CHAPTERS

a. The electric charge as special characteristic of a group of colloids

The most striking difference between the sols from Volume I and the sols which have been dealt with in Ch. VI, p. 153, is that in the former the colloid particles carry an electric charge but in the latter they do not. In this chapter systems will now be dealt with which in a certain sense form a bridge between Volume II and Volume I, since on the one hand their colloidal nature is due to the macromolecular character of the substances but on the other hand by the possession of electric charges they show similarities with systems in Volume I.

Before we proceed to discuss in detail the various aspects which the electric charge of the macromolecules involves, a general observation may first be made.

If the electrolyte character of a macromolecule is to be demonstrable, this must be dissolved in a polar solvent, for example water. Through the presence of this polar solvent, even apart from the charge of the particles, the interaction between macromolecule and solvent will be greater as a consequence of the stronger polar interactions (hydrogen bonds, etc.) than in non-polar media and it is certainly no accident that it is just the charged systems in a polar medium dealt with here, which have given the collective name "hydrophilic (or more generally lyophilic) colloids" to the whole group of colloids treated in this second Volume.

In this and in the succeeding chapters we shall deal with the various properties of the sols of macromolecular colloids in so far as they depend on the charge of the particles.

In the first place the charge itself demands our attention. It will appear that in most cases it is not produced by adsorption as in the systems of Volume I but is due to the dissociation of groups firmly attached to the macromolecule, such as COOH etc. The charge thus becomes dependent on the degree of dissociation and consequently on the pH. Nevertheless so-called indifferent electrolytes can also influence the charge and even make it reverse in sign (see Ch. IX, p. 259).

One of the most direct consequences of the presence of charge is the movement of the particles in an electric field. The electrophoresis of these macromolecules both in solution and in the adsorbed form will certainly be an important aid in the investigation.

The osmotic pressure, of the greatest importance with the uncharged macromolecules for the determination of the molecular weight, is that also in principle

for the charge-carrying macromolecules, but in this group the osmotic phenomenon is largely complicated by the occurrence of the DONNAN equilibrium which can frequently even quantitatively predominate over the osmotic action of the large molecules themselves. Since the DONNAN equilibrium is so closely connected with all double layer phenomena, it has been discussed extensively in Volume I. This discussion will not be repeated here.

The influence of the charge on the viscosity is a phenomenon very typical of this group of macromolecules. It is clear that in the presence of charges the skein will not assume its statistically most probable form, but will have a more rarified form through the interaction of the charged spots (mutual repulsion) or a more compact form if positive and negative charges are present simultaneously. As a result the viscosity, sensitive reagent for the skein form, is one of the most important aids in the investigation of systems with charged macromolecules.

Also the stability of the solutions of this group of charged macromolecules, characterised by the circumstances in which the solubility is just transgressed, has on account of the strong polar interactions a much more varied aspect than in the case of uncharged macromolecules. It appears that the charge can act both so as to raise and also so as to lower the solubility. The chapter on complex relations has this to thank for its particular significance.

b.   Shape and size of the kinetic units

In the previous subparagraph it has already been mentioned that the shape of the skeins is influenced by the electric charges and especially by their mutual interaction. If there are only charged spots of one sign on the macromolecule then the skein will be relatively rarified. If charged spots of both signs are present then the skein will just be more compact than corresponds to the most probable state. Since this remodelling of the skein depends on the interaction of the charged spots, it is dependent not only on the number of charged spots but also on the extension of the ion atmosphere around each charged spot. The shape of the skein will therefore also depend on the concentration and nature of the indifferent electrolytes in the solution since these determine the thickness of the ion atmosphere [1].

Furthermore it appears that among the charged macromolecules there also occurs a group of substances the solutions of which have a relatively low viscosity and the viscosity of which does not depend greatly on the electrolyte concentration. One must indeed assume for this and other reasons that one is not here dealing with skein shape (possibly modified) but that in these "corpuscular proteins" the macro-molecule is evidently rigid and — as is shown from the X-ray picture — folded in a very regular way. To what forces this formation is due is not yet known in full detail. In all probability cystine S—S linkages, salt bridges and hydrogen bonds play a part here.

The transition from the corpuscular type to the free skein shape is known from denaturation of proteins. It is quite in keeping with the above argument that this denaturation can be brought about by, among other things, extreme acid or alkaline

---

[1] c.f. two very recent communications on this subject given independently by W. KUHN and by J. J. HERMANS and J. TH. G. OVERBEEK, presented at the meeting on large molecules held at Liège (Belgium) in April 1948. To be published in *Bull. soc. chim. Belges*, 57 (1948).

media whereby all charged spots on the protein have the same sign and thus the expanding forces are as strong as possible, so strong that the forces binding the structure together can no longer resist them and the corpuscular structure ceases to exist.

In the following chapters attention will in the main be directed to the charged macromolecules of the statistical skein type not because the corpuscular proteins would be less important, quite the contrary, but because various excellent monographs on proteins [1] have appeared recently and they deal with this group of substances much more extensively than would be possible here.

c. Division of the colloids to be treated, according to the nature and sign of the charge of the ionogenic groups

The macromolecules with electrolyte character can be divided into three categories, to wit:

A. Colloids with acid character which carry only acid groups such as $-COO^{-}$, $-OSO_3^{-}$, $-OPO_3H^{-}$.

B. Colloids with basic character which carry only basic groups such as $-NH_3^{+}$, $-NH-C(NH_2)_2^{+}$.

C. Colloids with amphoteric character, which carry both sorts of groups.

The dissociation of the various groups can be more or less strong according to the pH (and the electrolyte content) of the medium. Thus for instance the carboxyl colloids of group A in strongly acid medium are practically uncharged.

The amphoteric colloids will at very low pH behave practically as "ionised colloids with basic character", at high pH just as "acid colloids [2]", while in the intermediate region positive and negative charges occur together and just equilibrate one another at the isoelectric point (I. E. P.).

With regard to the above division it may be remarked that types A and C both occur naturally and have been prepared synthetically. Pure representatives of group B, the basic colloids, are not known.

However we find substances which approximate closely to group B in the so-called basic proteins. Clupein, for example, consists to an extent of $^2/_3$ of arginine and contains no dibasic aminoacids. Along the whole chain there are thus 2 positive charged spots for each 3 peptide units and only one single COOH group at the end of the chain. In acid medium where the COOH group is not dissociated, cluplein behaves as a purely basic colloid while in less acid or alkaline medium also the positive character still greatly predominates.

If now we ask ourselves definitely what substances belong to the above groups, we are led automatically to a second kind of division, namely into:

Natural colloids: example gum arabic.

Modified natural colloids: examples, cellulose xanthogenate, pectate.

---

[1] CARL SCHMIDT, *The chemistry of the aminoacids and proteins*, THOMAS, Baltimore 1938.

E. J. COHN and J. T. EDSALL, *Proteins, amino acids, and peptides*, REINHOLD, New York 1943.

M. L. ANSON and J. T. EDSALL, *Advances in protein chemistry*, I, II, III, Academic Pi..s, New York 1944, 1945, 1947.

[2] This is however only correct to a first approximation with most of the proteins. In positive casein sols, for example, — casein contains phosphate groups — in the strongly acid region, where the dissociation of the carboxyl groups is entirely suppressed, negatively charged spots which originate in the phosphate groups are also present as well as the positive charge. Similarly in many proteins which contain arginine the total charge is also built up of negative and positive contributions in alkaline media on account of the strongly basic character of the guanidino group.

Wholly synthetic products: example polyacrylic acid.

We shall occupy ourselves in the following pages mainly with the natural and a few modified natural colloids, whereby, as we have already mentioned, the corpuscular proteins will be practically left out of consideration.

Since there are many among the natural substances in question the structure of which is not yet well or not generally known, and yet on the other hand it is desirable for a good understanding to have a schematic picture of the various substances involved in the investigation, we add here in the form of a list, schematic, highly simplified structural formulae of some important substances.

A. Colloids with COO^- as the carrier of the charge.
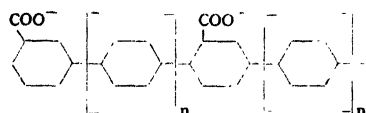
glycogen	$n = \infty$
soluble starch	$n =$ very large
oxidised starch	$n =$ large
gum arabic	$n =$ about 6
semen lini mucilage	$n =$ small
alginate	$n = 0$
pectate	$n = 0$
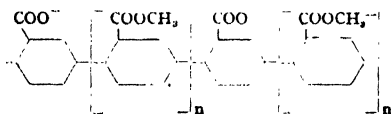
pectin	$n = \infty$
pectinate	$n > 0$
pectate	$n = 0$

B. Colloids with OSO_3^- as the carrier of the charge
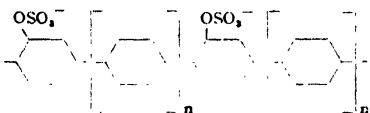
agar	$n =$ about 20
carrageen	$n = 0$

C. Amphoteric colloids with COO^- as the carrier of the negative charge

proteins
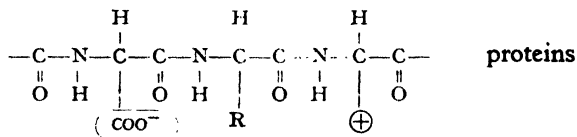
D. Amphoteric colloids with phosphate as negative group.

Thymus nucleic acid[1]. Scheme, see p. 188.

---

[1] See FREY-WYSSLING, *Submikroskopische Morphologie des Protoplasmas und seiner Derivate*, Borntraeger, Berlin (1938), p. 170.

In the scheme positive charges are drawn in the side chains to emphasise the amphoteric character[*]. The positive charges can be produced by dissociation of NH_2 groups which are permanently attached to rings of aromatic character. Consequently their basic character is very weak and dissociation only takes place in definitely acid medium (50% dissociation occurs in cytosine only at pH 4.2, in adenine at pH 3.7 and in guanine at pH 2.3). The I.E.P. lies very low (I.E.P. < 2). *All this results in nucleic acid behaving in the pH range 6—8 as a purely acid colloid in which phosphate groups are exclusively the carriers of the negative charge.*

At still higher values of the pH the negative charge increases still further through dissociation of aromatic OH groups in the side-chains (uracil and guanine: 50% ionisation at pH about 10) and at yet higher pH of OH groups in the sugar units (50% ionisation at pH > 13).

[*] For the dissociation behaviour of nucleic acid, nucleotides and nucleosides see: P. A. LEVENE and L. W. BASS, *Nucleic Acids*, Chemical Catalog Comp. New York (1931), p. 212, 280—286.

Yeast nucleic acid is probably similarly constructed. The sugar component is here ribose. It is a short chain consisting of only 4 ribose phosphate groups.

E. Association colloids with phosphate groups.

Further there is a group of substances which do not belong to the macromolecules but whose solutions behave as colloidal solutions and which will be used a few times as an example of colloids with phosphate groups as the carrier of the charge.

The substances in question are phosphatides such as lecithine, kephaline, etc.

The groups indicated by ($\ominus$) and ($\ominus$) are phosphate groups, the ($\oplus$) charge comes from choline in lecithine or colamine in kephaline), ($\bigcirc$) symbolises the phosphate group of the phosphatidic acid which is almost always present in the phosphatides and which can dissociate off one or two H· ions[1] according to the value of the pH. In the completely pure state the phosphatides would consist only of the symbols in brackets.

§ 2. NATURE OF THE ELECTRIC CHARGE OF MACROMOLECULAR COLLOIDS

a. Origin of the charge

How the charge of hydrophobic colloids comes into being has been discussed in detail in Volume I of this book. In this discussion the idea came to the fore that the carriers of the charge on the particles are in this case usually adsorbed ions. The degree of adsorption and with it the magnitude of the charge depends on the electrolyte concentration in the intermicellar liquid, in the first place on the concentration of the potential-determining ions and secondarily also on the concentration of the other (indifferent)[2] electrolytes.

---

[1] The scheme given is not intended to suggest a picture of the way in which the components are united into associated groups. These associations of groups are probably not linear but double sheeted planes. Compare on this point Chapter 14, *Association Colloids* and PALMER, K. J. and SCHMITT, F. O., *J. Cellular Comp. Physiol.*, 17 (1941) 385.

[2] In this and in the following chapters the term "indifferent electrolytes" will be frequently used for all those electrolytes that are not potential determining, or that do not give ions (H+ or OH−) which can react directly with the charged groups. The term "indifferent electrolyte" is preferred over the older term "neutral salt".

With the macromolecular colloids a different view on the origin of the electric charge appears however to be more appropriate. As already mentioned in § 1 c macromolecules can contain in their structure groups such as the carboxyl group, the sulphonic acid group, the amino group, etc., which by electrolytic dissociation can be the origin of the charge of these colloids.

Thus the proteins have carboxyl, amino, guanidino, imidazol groups, substances such as gum arabic and pectinic acid have carboxyl groups, agar ester sulphate groups, nucleates phosphate groups as carriers of the electric charge.

Many points of similarity can be found with the electrochemical behaviour of hydrophobic surfaces. In particular the dissociation of the above mentioned groups is greatly influenced by the pH, so that the H^+ ions (and the OH^- ions) can be considered as potential-determining ions.

The influence of indifferent electrolytes is also qualitatively the same as in hydrophobic systems. In both cases addition of the indifferent electrolyte increases the charge.

But if one wants to look for a quantitative explanation for the way in which the charge of macromolecular electrolytes depends on the composition of the medium, one should rather associate it with the theories of electrolytic dissociation such as also hold for small molecules than with the theory of the double layer.

b.   The influence of the pH and the composition of the solution on the charge of the macromolecule

From an electrometric titration with acid or alkali one can determine the charge of a macromolecular electrolyte as a function of the pH. It is true one encounters in the interpretation of the measurements all the difficulties [1] which are connected
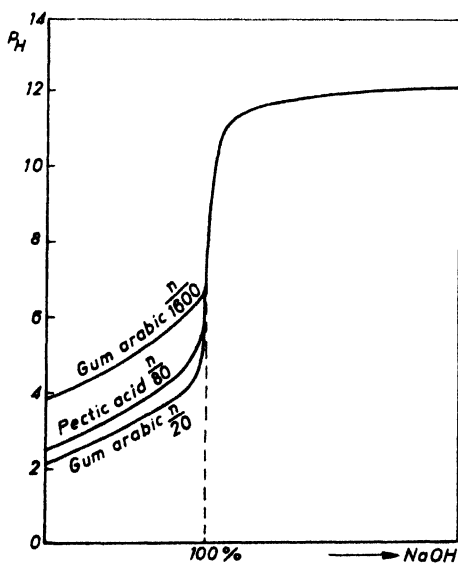
Fig. 1.   Titration curves of gum arabic and pectic acid.
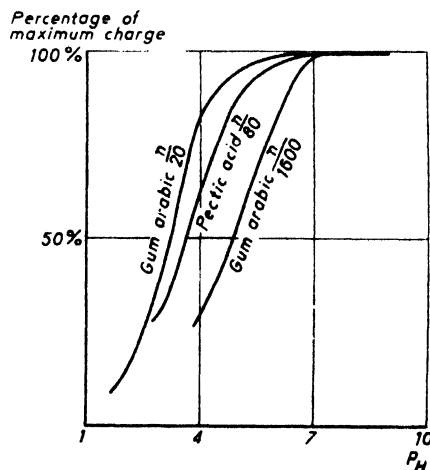
Fig. 2.   Relation between charge and pH of gum arabic and of pectic acid deduced from the titration curves.

[1] See Volume I, Ch. X, DONNAN *equilibria and sol-concentration effect.*

with the activity of a single ion (here the H^+ ion) not being unambiguously defined but nevertheless the application of the simple hypothesis, that the activity coefficients of the small ions are not influenced by the presence of the colloidal particles, has permitted a useful interpretation of the titration curves in various cases.

In the simplest cases we have to do with a macromolecule which contains a number of identical dissociable groups.

We give the titration curves of gum arabic[1] and pectinic acid[2] on page 189.

By means of the relation

$$Charge = added\ NaOH + conc.\ H^+ - conc.\ OH^-$$

in which C_{H^+} and C_{OH^-} are calculated from the pH assuming that the activity coefficients are 1 (or a known constant < 1 in a medium containing a salt), one can readily deduce curves, from the titration curves, which give the connection between charge and pH (see Fig. 2).

If the carboxyl groups are all linked in the same way in the macromolecule and are situated so far apart that their interaction can be neglected, the titration curves ought to be identical with those of a monobasic weak acid. This appears however not to be the case. If one determines the dissociation exponent pK at various points of the curve

$$pK = pH - \log \frac{[A^-]}{[HA]}$$

in which $[A^-] = [H^+] + [Na^+] - [OH^-]$ and $[HA] =$ gross concentration of colloid $-[A^-]$ then pK is found to rise with increasing degree of neutralisation, in pectin for example from 2.7 at the beginning of the titration to 4.0 close to the equivalence point.

One could interpret this by assuming that strong acid groups first react, then weak acids and ascribe this difference to a difference in binding of the carboxylic groups.

An explanation more in accord with our knowledge of the structure of these substances starts from the assumption that the intrinsic dissociation constants of all the acid groups is the same, but that dissociation is more difficult the higher the charge of the macromolecule. The negative charges already present make, by a simple electrostatic interaction, a further dissociation more difficult. As the interaction between two charges in an electrolytic medium is weakened by an increase of the ionic strength, it is to be expected that the change in pK becomes less and less important, the higher the ionic strength. A quantitative relation between pK, the charge of the large molecule and the ionic strength has been given very recently independently by KUHN and by OVERBEEK[3].

In the titration of proteins which through the occurrence of different kinds of dissociating groups are themselves more complicated than the cases sketched above,

[1] A. W. THOMAS and H. A. MURRAY, *J. Phys. Chem.*, 32 (1928) 676; D. BRIGGS, *J. Phys. Chem.*, 38 (1934) 867.

[2] J. BONNER, *Proc. Acad. Sci. Amsterdam*, 38 (1935) 346.

[3] W. KUHN and coworkers, J. TH. G. OVERBEEK, *Meeting on large molecules held at Liège (Belgium)* in April 1948. To be published in *Bull. soc. chim. Belges*, 57 (1948)

the determination of dissociation constants has been carried out with more care and on a greater scale.

With certain modifications due to the globular character of the proteins, in contradistinction to the more open skein structure of the acids mentioned above, the same interpretation can be used here [1].

As an example, in Fig. 3 we give the charge — pH curve [2] deduced from a titration curve of carboxyhemoglobin.

This curve clearly exhibits a number of different steps which obviously originate in different types of charge-carrying groups. Other proteins also show a similar variation of charge with pH.

It has been possible to analyse these titration curves and attribute the various steps to the carboxyl groups of glutamic acid and aspartic acid (to about pH = 5), the imidazol group of the histidine (pH=5—9) NH_2 groups of lysine (pH=9—11) and the guanidino group of arginine (pH ∞ 12) and perhaps also the phenolic OH of tyrosine at very high values of the pH.

For various proteins these data deduced from the titrations agree well with other analytical data such as amino acid composition, amino nitrogen content etc. [3].

Fig. 3. Relation between charge and pH of carboxyhemoglobin.
———— practically salt-free.
- - - - - - in 1 M. NaCl.

On addition of salts the titration curves show a steeper course as can be seen in Fig. 3 where the full line holds for a salt-poor medium and the dotted line for a medium which is 1 M in NaCl. This steeper course is readily explained if one remembers that in the salt-rich medium the interaction between the various charge - carrying groups becomes smaller through the compression of the ion atmospheres and consequently therefore the dissociation proceeds more easily at the same pH.

The dissociation behaviour of the phosphatides discussed on page 188 is also interesting. Since choline is a strong base and phosphoric acid a strong acid, the I.E.P. of pure lecithin lies very close to the neutral point, and in a broad pH range around the I.E.P. the lecithin is present almost exclusively as amphoion (with net charge =0). (Fig. 4a). In this pH range the + and — charge compensate one another almost completely (Fig. 4b full curve).

A small admixture of phosphatidic acid in lecithin (as supposed in the scheme on p. 188) must then result in the full curve in Fig. 4b being shifted somewhat down-

[1] K. LINDERSTRØM-LANG, Compt. rend. trav. lab. Carlsberg, 15 (1924) 7.
R. K. CANNAN, A. KIBRICK, and A. H. PALMER, Ann. New York Acad. Sci., 41 (1941) 243.
[2] E. J. COHN, A. A. GREEN, M. H. BLANCHARD, J. Am. Chem. Soc., 59 (1937) 509.
[3] See for example β-lactoglobulin, R. K. CANNAN, A. H. PALMER, A. C. KIBRICK, J. Biol. Chem., 142 (1942) 803, and E. BRAND, L. J. SAIDEL, W. H. GOLDWATER, B. KASSEL, F. J. RYAN, J. Am Chem. Soc., 67 (1945) 1524.

wards as a whole as a result of which the I.E.P. is displaced very rapidly to lower pH's. Since the phosphate group of the phosphatidic acid possesses two hydrogen atoms whicn can be dissociated off, it may be anticipated that with this shift downwards an additional S-shaped course in the curve (as indicated in the dotted curve) will appear at the same time in the neighbourhood of the pK value corresponding to the second dissociation constant of the phosphate group.

Lecithin preparations produced in practice (without very rigorous precautions) do indeed possess I.E.P. which are very much lower than the theoretical I.E.P. (p. 294 Ch. IX § 2l). As a result ot the facts mentioned above they behave in the neighbourhood of the neutral point as colloids with acidic character, in which the phosphate group is the carrier of the negative charge.

Similar considerations hold with regard to kephalin, only as a result of the weaker basic character of colamine (compared with choline) the I.E.P. here lies at a somewhat lower pH in the pure substance than is the case with pure lecithin. Here also admixture of a small amount of phosphatidic acid will result in a considerable lowering of the I. E. P.

Fig. 4a. The dissociation of the phosphate group and of the choline group occurring in lecithin.
Fig. 4b. ———— The charge of theoretical lecithin as a function of the pH
- - - - - The charge of lecithin, containing some phosphatidic acid

It is also worth while pointing out that with macromolecular electrolytes the charge bears always a statistical character. Thus in the case of hemoglobin in which the isoelectric point is situated in the dissociation region of the imidazol groups (33 per molecule) one can calculate that at the isoelectric point (pH = 6.4) only about 20 percent of all the molecules are really isoelectrical at any given instant and that molecules with a charge of + 3 or — 3 units are certainly not rare. Table I gives an idea of the charge distribution for the various molecules assuming that of the 33 imidazol residues 13 have a dissociation exponent of 5.7 and 20 a dissociation exponent of 7.5, which fits as closely as possible with the titration curve of Fig. 3.

For a more detailed discussion of the topic dealt with above reference may be made to the book of Cohn and Edsall [1].

Summarising one can say that the charge of macromolecular electrolytes is above all governed by the acid (or base) strength of the dissociating groups and by the pH of the medium, while the salt concentration also exerts a secondary influence.

In the part of this chapter in which the viscosity is treated and in Chapter IX (on reversal of charge), we shall see that besides this dissociation of the specific groups

---

[1] E. J. Cohn, J. T. Edsall, *Proteins, amino acids and peptides*, Reinhold, New York 1943. See especially Ch. 20, *Proteins as acids and bases.*

TABLE 1

CHARGE DISTRIBUTION OVER THE VARIOUS MOLECULES OF HEMOGLOBIN
AT THE ISOELECTRIC POINT pH = 6.4

Charged imidazol residues	Uncharged imidazol residues	Charge on the molecule	% molecules which carry the charge given in column 3
25	8	+ 4	1.26
24	9	+ 3	3.90
23	10	+ 2	9.42
22	11	+ 1	16.98
21	12	0	22.40
20	13	− 1	21.20
19	14	− 2	14.30
18	15	− 3	6.99
17	16	− 4	2.65
16	17	− 5	0.71

adsorption of other ions must also play a part in the establishment of the charge. Although this effect may not always be evident in the amphoteric proteins [1], the rôle of adsorption is brought out more clearly in the reversal of charge phenomena in the case of macromolecules with exclusively negative groups. In fact by suppression of the dissociation we should at most be able to obtain discharge while in reality we observe that the charge from being negative can become positive. Here we have to do with a direct analogue of reversal of charge in hydrophobic colloids, which could also not be understood from the theory of the diffuse double layer but wl. re in addition specific adsorption of ions (or charged complexes) had to be brought in for the explanation.

§ 3. THE ELECTRIC CHARGE OF MACROMOLECULES AS DETER-MINED BY ELECTROPHORESIS

a. The interpretation of the electrophoresis of macromolecules

In principle one can also establish the charge of a macromolecule from electrophoresis measurements. This method has even definite advantages over the analytical method dealt with in the previous §. Indeed in the titration one only determines that part of the charge which is based on binding of H^+ or OH^- ions. In electrophoresis one also takes into account charge of a different origin, that is to say, produced by adsorption of other than H^+ or OH^- ions. Furthermore it is frequently difficult with the titration method to fix the charge-zero while this can be done very easily with electrophoresis.

On the other hand however the quantitative interpretation of the electrophoresis still raises various difficulties, whereby the conclusions will again become less certain. For a detailed discussion of electrophoresis, also as concerns the methodics, reference may be made to Chapter V of Volume I. We shall however deal briefly here with the interpretation of electrophoresis.

---

[1] On this point see however § 3b, the comparison of electrophoresis and titration.

To a first approximation one can calculate the ζ potential from the electrophoretic velocity ($E.V.$) with the aid of SMOLUCHOWSKI'S relation

$$\zeta = \frac{4 \pi \eta}{\varepsilon} \cdot v \tag{1}$$

or with the refinements given by HENRY[1] and others.

This relation holds for compact particles whose dimensions are large compared with the thickness of the double layer. For smaller particles various corrections must be applied of which especially the one for the relaxation of the ion cloud still contains difficulties[2].

Next one must calculate the charge from the ζ potential and this brings two new difficulties with it. As is well known the ζ potential is the potential on the "slipping-plane" between particles and liquid. In general a layer of liquid of one or more molecules thick will stick hydrodynamically to the particle and quite a portion of the ions from the outer coating of the double-layer may be present in this layer. As a result of this the charge calculated from the ζ potential is always lower than the charge of the particle itself and this lowering can be important especially at higher electrolyte concentrations. A second difficulty is formed by the fact that calculation of the charge from ζ calls for accurate knowledge of the shape and size of the particle and this is not always easy to obtain. Determinations of the molecular weight and the velocity of diffusion leave two interpretations open if the diffusion is slower than corresponds to a compact sphere. Either one assumes that the particle is spherical but enlarged by hydration or that there is no hydration but the particle has an elongated form. And the two hypotheses lead to different values of the charge for the same value of ζ[3].

As a general conclusion one may state that, provided the ζ potential, and thus the electrophoretic velocity, is small ($\zeta < 25\ mV$, $v < 1\ \mu\ cm/Vsec$) the charge of the particles is directly proportional to ζ and therefore also to v. The exact indication of the proportionality factor (which also depends on the electrolyte concentrations of the system) still presents difficulties.

The above statements hold for the electrophoresis of compact particles. These considerations cannot be applied to a charged skein molecule without further discussion. Apart from the motion of the particle with respect to the surrounding liquid, the motion of the liquid through the meshes of the skein molecule has certainly to be taken into account. No theory has yet been given for the electrophoresis of clewed molecules. The only thing which one may expect with high probability here is that, again for the region of small velocities charge and electrophoretic velocity are proportional to one another and that electrophoretic stand-still thus signifies that the macromolecule carries no charge. But in this case we are even worse informed regarding the proportionality factors between charge and electrophoretic velocity than in the case of compact particles.

A much used expedient in the electrophoretic investigation of macromolecules is the determination of the $E.V.$ not of free molecules but of rather coarse particles

[1] D. C. HENRY, *Proc. Roy. Soc. London*, 133 (1931) 106.
[2] See Chapter V, part I and J. TH. G. OVERBEEK. *On the interpretation of the electrophoretic velocity*, in *Advances in Colloid Science III*, Interscience New York, in the press.
[3] See H. A. ABRAMSON, L. S. MOYER, and M. H. GORIN, *Electrophoresis of proteins*, New York 1942, p. 151 *et seq.*

(quartz, carbon, oil drops etc.) which are covered to saturation with macromolecules on their surface. The *E.V.* of these coarse particles can then be determined very readily by microelectrophoretic methods.

Since now we are dealing with very large particles, SMOLUCHOWSKI's electrophoresis equation (1) can be applied with confidence and the ζ potential is therefore calculable. From ζ the charge per cm² can then be calculated if the electrolyte concentration is known but to find the charge per macromolecule from this, the density of covering of the surface must be known and that is in general not known. Thus with this method also we do not get any further than to a relative measure for the charge of the macromolecules in which an influence of the absorbing particle on the electrochemical properties of the absorbed substance can possibly also occur as a complication.

b. Some examples

In the first place we give some comparisons between the electrophoretic velocity of corpuscular proteins in solution and of the same proteins absorbed on various particles under the same conditions as regards pH and salt concentration. It appears that in various cases these two *E.V.* are identical (serum albumin, Fig. 5, pseudo-globulin, Fig. 6) and that in a case such as ovalbumin (Fig. 7) where a definite difference exists, this difference is nevertheless small and can practically be described as a shift in the direction of denaturation which results in the *E.V.* of the absorbed oval-bumin resembling more that of the dena-tured protein.
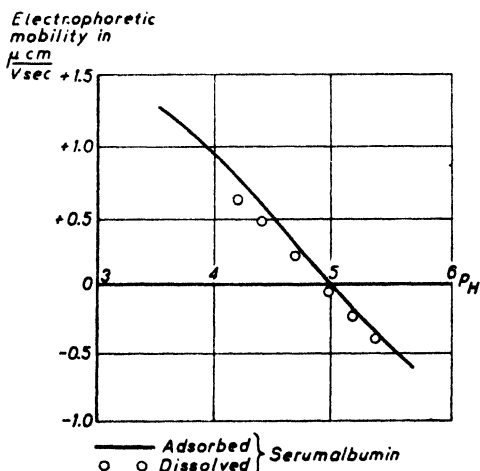
Fig. 5. Electrophoretic velocity of dissolved and adsorbed serum albumin.
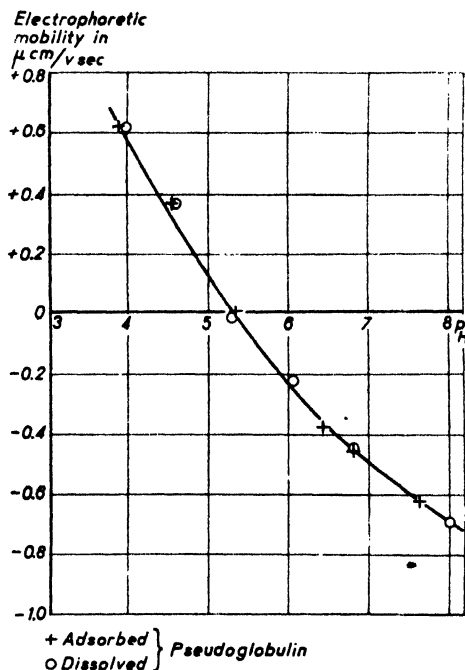
Fig. 6. Electrophoretic velocity of dissolved and adsorbed pseudoglobulin.

Although fundamentally this identity or great similarity is unexpected and not at all explained, one can nevertheless conclude from it that one obtains valuable information by electrophoretic investigations on adsorbed macromolecules even on

the behaviour of these substances in solution, which is of great importance since the measurements with the adsorption method are so much simpler.
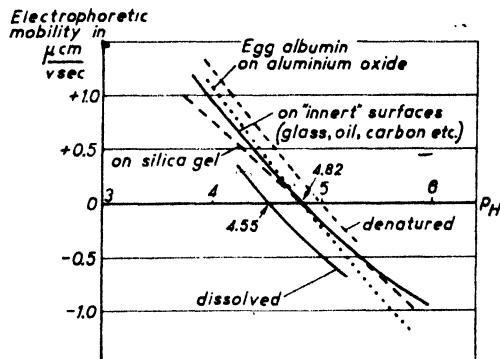
Fig. 7. Electrophoretic velocity of ovalbumin, in the dissolved state, adsorbed on particles of acidic (silica gel), basic (Al₂O₃) and inert (oil, carbon etc.) nature and in the denatured state (surface denatured).

Electrophoretic measurements by this adsorption method on many biocolloids of the skein type will be dealt with in Chapter IX (p. 259). Unfortunately a comparison with the free, non-adsorbed, substances is here entirely missing.

Next, it is worth while comparing the data from titration curves with those from electrophoresis. Here again the best investigations have been made on corpuscular proteins. A beautiful example is to be found in the work of CANNAN, KIBRICK and PALMER [1] on the titration and that of LONGSWORTH [2] on the electrophoresis of ovalbumin.

Electrophoresis measurements are indicated in Fig. 8 by circles. The titration curve is made to fit as well as possible with the electrophoresis measurements by assuming the same I.E.P. for the titration as was found in electrophoresis and by suitably choosing the scale of the titration. It is seen that a good proportionality then exists between E.V. and titrated charge. Making use of the molecular weight of ovalbumin (45 000) one can then calculate that the following empirical relation exists between the charge Q of one molecule of ovalbumin in terms of the electron charge and the E.V. (v) in μ cm/Vsec:

$$Q = 36 \, v$$

Application of HENRY's electrophoresis formula [3] treated in Volume I leads however to the relation:

$$Q = 22 \, v$$

while a somewhat refined calculation of ABRAMSON, MOYER, and GORIN [4] also leads to a similar value. The electrophoretic

Fig. 8. Comparison of the electrophoretic velocity and the charge as deduced from titration.
o o o o o measured electrophoretic velocity
———— titrated charge
The I.E.P. of the titrated charge has been arbitrarily made to coincide with the electrophoretic I.E.P. The ordinate scales were so chosen that the electrophoretic velocities and the titration approximately coincide.

[1] R. K. CANNAN, A. KIBRICK, and A. H. PALMER, Ann. New York Acad. Sci., 41 (1941) 243.
[2] L. G. LONGSWORTH, ibid., p. 267.
[3] For details see the article by LONGSWORTH quoted above.
[4] H. A. ABRAMSON, L. S. MOYER, and M. H. GORIN, Electrophoresis of proteins, New York 1942, p. 152 et seq.

velocities, which play a part here, are still so small that one can indeed neglect the relaxation correction.

Thus in the best investigated case in this line there is an important difference between the electrophoretic charge and the titrated charge and of such a kind that the titrated charge is appreciably larger.

An explanation which is satisfactory in all respects has not yet been given for this. Attempts in this direction have been made by various workers[1], in which the suggestion is mainly that of an attachment of other ions than H^+ and OH^- (for example Cl^- ions) to the protein. Qualitatively another good explanation of this anomaly can be given by taking into account an adsorbed water-layer (see Volume I, Ch. V). Indeed at the fairly high electrolyte concentrations used in this case (0.1 n) a by no means negligible portion of the counter ions may already have been taken up in this layer.

Another pretty example[2] of the application of electrophoretic velocities is the comparison of gelatin with "desaminised gelatin". It can clearly be seen in Fig. 9 that desamination amounts to the removal of

Fig. 9.   Electrophoretic velocity of normal and desaminised gelatin.

positive charged spots, whereby the whole electrophoresis curve shifts downwards.

§ 4. FORMER CONCEPTIONS ON THE NATURE OF THE KINETIC UNITS IN HYDROPHILIC SOLS OF THE HIGH VISCOUS TYPE

KRUYT, BUNGENBERG DE JONG and coworkers[3], developed some twenty years ago a theory of hydrophilic colloids, which may be characterised briefly as follows: Their sols must be regarded as two-phase systems, in the same sense as this is convenient in discussing hydrophobic sols. The kinetic units of their sols are not different from those in lyophobic sols as regards the occurrence of a capillary electric charge (characterised by a ζ potential) at the boundary of the particles. They only differ in having a strong solvation, which acts as an extra stability factor. This solvation may be regarded either as a thick shell of hydration around the solid particle, or as a less thick shell of hydration water surrounding the swollen particle, but always with the characteristic feature, that a sharp periphery of the hydration shell is absent.

[1] H. A. ABRAMSON, L. S. MOYER, and M. H. GORIN, l.c., p. 158. L. J. LONGSWORTH, l.c.
J. STEINHARDT, Ann. New York Acad. Sci., 41 (1941) 287.
[2] H. A. ABRAMSON, J. Gen. Physiol., 15 (1932) 575.
[3] H. R. KRUYT and H. G. BUNGENBERG DE JONG, Kolloidchem. Beih., 28 (1928) 1 and other publications with coworkers in the Kolloidchem. Beihefte under the title Zur Kenntnis der lyophilen Kolloide.

Only such a "diffuse solvation shell" could account for the solvation being a stability factor.

"Hydrophobic sols" have only one stability factor: the capillary electric charge, "hydrophilic sols" however have two stability factors: capillary electric charge and hydration. Flocculation of the latter will only occur if both factors are sufficiently suppressed. This condition can be reached in steps along two ways: see Fig. 10.
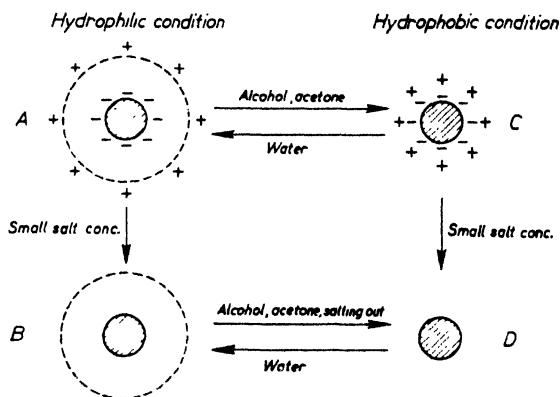
Fig. 10. Change of state of the kinetic units of hydrophilic colloids according to earlier ideas.
A. In pure water the kinetic unit is characterised by the presence of a double layer and a diffuse hydration (here represented as an external shell, but can also be taken up for a great part inside the particle). B. After addition of a small amount of salt: the double layer (not shown) has receded inside the hydration shell. C. and D. are similar to A and B as regards charge but do not possess the diffuse hydration shell. A possesses two, B and C one and D no stability factors.

A. by first removing the charge and afterwards the hydration, as by adding $MgSO_4$ in increasing quantities to an agar sol. In small concentrations the capillary electric charge is suppressed, in much higher concentrations dehydration sets in, the colloid is "salted out".

B. by first dehydrating and then removing the charge, as can be realised with sufficiently diluted agar sols, with alcohol and afterwards adding a little of an indifferent salt.

This stability theory, which was also extended to amphoteric colloids — at first proved a useful guide in experimental work (e.g., also in the first stages of the investigations on coacervation, see p. 243, Chapter VIII, § 3). But gradually facts were met with which are not compatible with it. Its usefulness became less and less and nowadays we may say it has only historical value.

Indeed at the present time there seems no longer to be any real need for a "stability" theory of hydrophilic sols, as we now regard the colloid substance as truly dissolved, and its electrical properties as caused by ionisation of ionogenic groups in the dissolved macromolecules. Flocculations or coacervations are now regarded as transgressions of solubility. Many of the facts, the above stability theory seemed to explain have received other explanations. See for instance for the opalescent sols with lyophobic character obtained by adding alcohol to the sol Chapter VIII § 1c (p. 234), and for the flocculation or coacervation with alcohol + indifferent salt Chapter X § 3f (p. 396).

The viscosimetric data, upon which for a greater part the now discarded stability theory was based, remain, and they need a reinterpretation from the macromolecular point of view.

Before doing so, we give in the next subparagraph first a survey of these data (and their bearing to the former conceptions).

§ 5. CHANGES IN $(\eta_s - \eta_o)/\eta_o$ AT CONSTANT COLLOID CONCENTRATION, CAUSED BY ALTERING THE COMPOSITION OF THE SOLVENT

In all theories which try to connect viscosity of sols with the state or structure of its kinetic units, this connection is always made through relative viscosity η_s/η_o or relative viscosity increase $\dfrac{\eta_s - \eta_o}{\eta_o}$ now frequently symbolized by η_{spec}, in which η_s is the viscosity of the sol and η_o the viscosity of the dispersion medium (in older conceptions of hydrophilic colloids) or of the solvent (modern conceptions).

KRUYT, BUNGENBERG DE JONG and coworkers based their stability theory on the changes that the relative viscosity of hydrophilic sols shows if at constant colloid concentration (and constant temperature) the composition of the "dispersion medium" is systematically altered. In using the term relative viscosity it will be clear, that it has here a different meaning from elswhere in physical chemistry.

Instead of comparing the measured viscosities with that of a single constant calibration liquid, η_s/η_o means here the values to be calculated with reference to the viscosity of the dispersion medium (at the same temperature), which medium each time has a different composition. Thus below in Fig. 12 the influence of increasing concentrations of alcohol on $\dfrac{\eta_s - \eta_o}{\eta_o}$ of a dilute agar sol is given. This figure is obtained by calculation from the two sets of viscosity determinations represented in Fig. 11, viz., from the viscosities of a dilute agar sol containing increasing amounts of alcohol (upper curve), and from the viscosities of the corresponding alcohol water mixtures (lower curve). In these two sets of determinations water is taken as calibration liquid. By division of the corresponding viscosity values a series of relative viscosities in the sense as above indicated is obtained, and by subtracting

Fig. 11. Viscosity measurements. (45° C) required for the calculation of the curve of Fig. 12.
Lower curve: alcohol-water mixtures.
Upper curve: alcohol-water mixtures which contain 0.134 g agar per 100 ml.
Ordinates: Relative viscosities with water at 45° as calibration liquid.
Abscissae: Alcohol content in weight percents.

unity from them, the values of $\dfrac{\eta_s - \eta_o}{\eta_o}$ used as ordinates in Fig. 12 are obtained.

Al further viscosity changes to be discussed below refer to relative viscosities in the above sense; we keep further to a constant low colloid concentration and a constant temperature and only vary the composition of the dispersion medium.

The now following survey falls into two sections, which relate to obviously

different causes for the changes in relative viscosity. We have labelled them provisionally "*Solvation*" and "*Electric Charge*", thus following more or less the interpretations suggested in the stability-theory of § 4 (p. 197). A reinterpretation from the macromolecular point of view will be postponed till later (p. 209 § 6).

A: "*Solvation*".

For an interpretation of viscosity data of sols obeying POISEUILLE's law KRUYT and BUNGENBERG DE JONG took EINSTEIN's equation

$$\eta_s = \eta_0\,(1 + 2.5\,\varphi) \quad \text{or} \quad \frac{\eta_s - \eta_0}{\eta_0} = 2.5\,\varphi$$

as a starting point. In this formula φ is the ratio of the volume of the dispersed phase and the volume of the whole sol.

Determinations performed with dilute sols (the formula indeed has been deduced for values of φ not exceeding a few per cent (by volume) of many colloids (e. g., agar, gum arabic, amylum solubile, gelatin and others) have shown that frequently the value of φ as calculated from the above formula with the factor 2.5, appears to be many times (e.g., in some cases more than $100 \times$) larger than could be expected from the volume of the dispersed substance in the dry state. Now the formula is valid for rigid particles of spherical shape only. With other shapes of particles we must write $(\eta_s - \eta_0)/\eta_0 = k\,\varphi$ (SMOLUCHOW- SKI [1]), where $k > 2.5$.

In looking for an interpretation of the much too great values of φ it is consequently uncertain whether this should be attributed to the non-spherical shape, or to a swelling (solvation) of the particles.

So long as only one dispersion medium is used, there is no possibility of distinguishing between these two cases for a given temperature and given concentration of the dispersed substance.

If the too high values for φ originated from the first cause (rigid non-spherical particles), according to the EINSTEIN formula, determinations performed with various dispersion media (variation of η_0), provided both temperature and colloid concentration were kept constant, would always lead to the same value of $(\eta_s - \eta_0)/\eta_0$.

Experimentally this appears not be the case. Here then we find a starting point
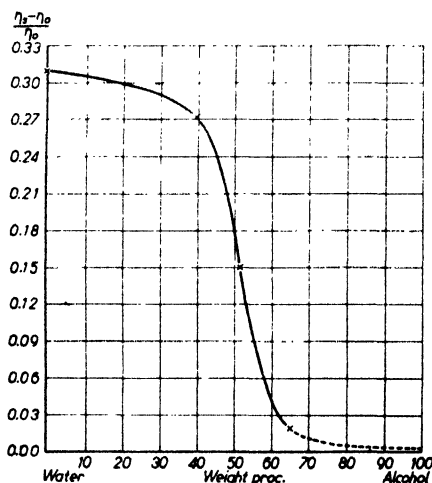
Fig. 12. Influence of alcohol on $(\eta_s - \eta_0)/\eta_0$ of a 0.134% agar sol (45° C).

The ordinates are obtained by dividing the values η_s/η_w of the upper curve in Fig. 11 by those of the lower curve (at the same alcohol content) and then subtracting one from the values obtained for η_s/η_0 (relative viscosities in the sense described in the text).

Dotted part of the curve: probable further course of the curve, assuming that the latter at 100% alcohol ends at the value 0.002 which follows from EINSTEIN's formula (with factor 2.5) for dry agar (density about 1.5).

[1] M. VON SMOLUCHOWSKI, *Kolloid-Z.*, 18 (1916) 190.

for experimental work. By systematically changing the dispersion medium, results may be obtained which will show us changes in solvation (or in solvation $\vdash$ amount of occlusion liquid).

For a number of negatively charged hydrophilic sols curves of the type represented in Fig. 12 are obtained if alcohol or acetone is gradually substituted for water (the diagram reproduced refers to an agar sol in water-alcohol mixtures). It is seen that in the concentration range 40—60% alcohol there is a strong decrease of $(\eta_s - \eta_o)/\eta_o$. At the same time the stability changes in a conspicuous way. Whereas the sol in aqueous surroundings appears stable against the addition of salts such as KCl or BaCl$_2$, this is no longer the case after the descent of the curve. It is now exceedingly sensitive to the addition of electrolytes, minute quantities of which will bring about flocculation or coacervation. It has thus obtained the character of a lyophobic sol (see stability theory p. 198, § 4).

The decrease of $(\eta_s - \eta_o)/\eta_o$ from the point of view of EINSTEIN's equation must be interpreted as a process of *desolvation* (the much too high values of η ascribed to solvation, being reduced to values nearer the calculated ones).

Besides the type mentioned there is a second type of curve [1] which is obtained upon addition of tannin, (sols of agar, amylum, gelatin and in the case of proteins also upon addition of lower phenols). Some instances have been represented in Fig. 13a (agar sol at 40° C and 50° C upon addition of tannin) and Fig. 13b (gelatin sol upon addition of resorcinol).
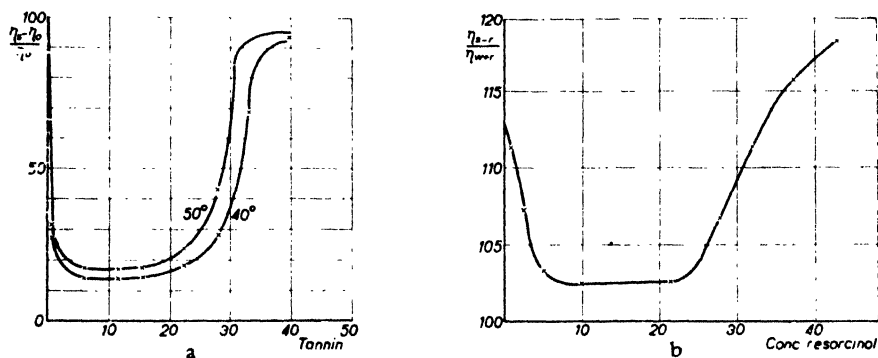
Fig. 13   Influence of tannin on a dilute agar sol (a, 40° and 50° C) and of resorcinol on a dilute isoelectric gelatin sol (b, 42° C).

Ordinates: $(\eta_s - \eta_o)/\eta_o$ (in % of the initial value), or $\eta_s/\eta_o(\eta_o$ is in this case always the viscosity of the corresponding tannin - water or resorcinol-water mixture).

Abscissae: gr. tannin or resorcinol per 100 ml.

The curves drop fairly steeply at relatively small concentrations, rise afterwards in the region of large concentrations. With gelatin the curve rises above the initial value. The clearly noticeable lowering of the pH as a consequence of the very high resorcinol concentration may contribute to this effect.

The effect produced by tannin and polyphenols is fairly strongly dependent on the temperature; it decreases as the temperature rises. This is also evident from the relative positions of the two curves in a.

[1] H. G. BUNGENBERG DE JONG, *Rec. trav. chim.*, **42** (1923) 437; **43** (1924) 36; **46** (1927) 727; **84** (1929) 494.

In this case we first have a strong decrease of $(\eta_s - \eta_o)/\eta_o$ for relatively small concentrations of the added substance and afterwards an increase at much higher concentration.

As appears also from the changes in stability we must conclude that a desolvation
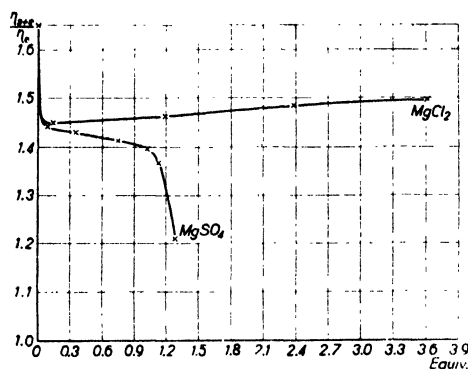
Fig. 14. Influence of $MgCl_2$ and $MgSO_4$ on the relative viscosity of a 0.14% agar sol (50° C).

Ordinates: relative viscosities (relative to the corresponding salt solution without agar).

Abscissae: salt concentration in equiv. per 1.

After a considerable fall at the small concentrations, which is characteristic of all electrolytes, the relative viscosity changes but little for $MgCl_2$ (increases), but with $MgSO_4$ a sharp drop begins at about 1 N. The further course of the $MgSO_4$ curve downwards could not be measured on account of salting-out. One may expect a course in this case such as that for alcohol at high concentrations in Fig. 12.
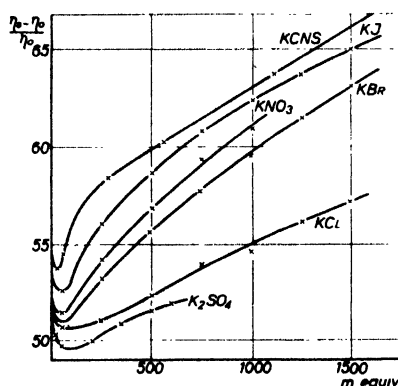
Fig. 15. Influence of a number of K salts on $(\eta_s - \eta_o)/\eta_o$ in a 0.5% sol of amylum solubile (MERCK) at 25° C.

Ordinates: $(\eta_s - \eta_o)/\eta_o$ (in % of the initial value), in which η_o is throughout the viscosity of the corresponding salt solution at the same temperature.

Abscissae: salt concentration in milli equiv. per 1.

The curves begin at $(\eta_s - \eta_o)/\eta_o = 100$ and drop very rapidly in the first 10 m. eq. per 1.

The text discusses the significance of the relative positions of the curves in the region of higher concentrations (10—1500 m. eq. per 1) which is that of the so-called lyotropic series of anions.

process takes place at the lower concentrations, whereas at a higher concentration the solvation increases again considerably.

Not only non-electrolytes (alcohol, acetone, tannin, resorcinol etc.) but also electrolytes may alter the "degree of solvation". An example of curves obtained in this way has been given in Fig. 14, representing the effect of $MgSO_4$ and $MgCl_2$ on the relative viscosity of a diluted agar sol.

After a decrease at very small electrolyte concentrations, the meaning of which will be discussed in B (p. 203), further addition of $MgCl_2$ changes the specific viscosity very little, $MgSO_4$ on the other hand causes a second decrease at higher concentration. Now $MgCl_2$, even in the highest concentrations does not flocculate the agar sol; salting out with $MgSO_4$, however, occurs exactly at the point where the curve could not be measured further downwards. Hence from the point of view of the EINSTEIN equation: this second decrease in specific viscosity is due to a desolvation process. $MgCl_2$, instead, seems to increase the solvation slightly.

With other sols electrolytes may give similar effects. After a strong decrease at very low concentrations, to be discussed below, at higher concentrations $(\eta_s - \eta_o)/\eta_o$ may increase, which must be interpreted as increase of solvation. Fig. 15 gives an example, referring to amylum solubile[1] and a number of potassium salts. The interpretation, that these changes in spec. viscosity are really related to changes in solvation seems fortified by the order of the curves, which is that of the so-called lyotropic series of the anions: $CNS > J > NO_3 > Br > Cl > SO_4$

B: "Electric Charge" (Electroviscous Effect)

In Figs. 14 and 15 we found a peculiar descent of $(\eta_s - \eta_o)/\eta_o$ at very low electrolyte concentrations, which descent can hardly be interpreted as a variation of the degree of solvation. Indeed, if the effect of various types of electrolytes is measured, there is no indication of the appearance of a lyotropic order. SCHULZE-HARDY's rule on the other hand, comes conspicuously to the front: the curves are arranged in narrow bundles, for each of which the valency of one of the ions is characteristic. See Fig. 16, which, for the agar sol[2], shows the results obtained within the concentration range 0—4 m.eq. p. l. It appears, that the fourteen investigated salts give four bundles of curves. The upper bundle comprises KCl, NaCl, LiCl, NH₄Cl, KCNS, K₂SO₄ and K₄Fe(CN)₆, that are salts having in common a *monovalent cation*.

The second bundle unites salts having a *divalent cation*, viz., BaCl₂, SrCl₂, MgSO₄, CdSO₄. The third bundle contains salts with a *trivalent cation* viz., La(NO₃)₃ and Co(NH₃)₆Cl₃, whereas the fourth bundle, characterised by a *tetravalent cation*, is represented here only by a complex platinum salt [Pt(en)₃](NO₃)₄ where (en) stands for ethylenediamine.

The results obtained strongly recall the influence of indifferent salts on a negatively charged "lyophobic" sol, where also valency of the cations is of primary importance in lowering the capillary electric potential (see Volume I). As in the electric field agar moves to the anode, the kinetic units in
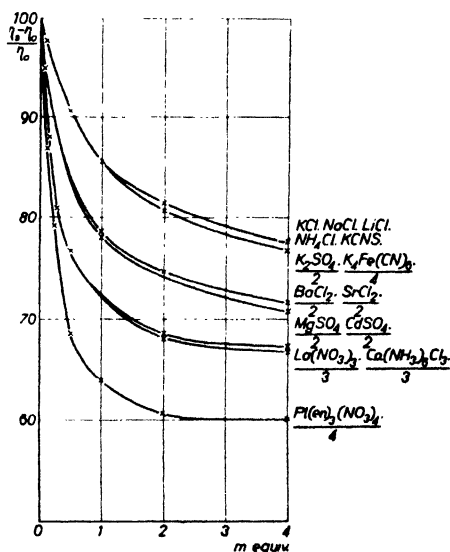
Fig. 16. Influence of salts on $(\eta_s - \eta_o)/\eta_o$ of a 0.14% agar sol (50° C) in the region of small concentrations (Neutralisation of the electroviscous effect).

Ordinates: $(\eta_s - \eta_o)/\eta_o$ in % of this value without added salt.

Abscissae: salt concentration in m. eq. p. l.

The salts arrange themselves in four bundles each characterised by the valency of the cation.

With the complex Pt salt opalescence appears along the lower portion of the curve (see § 10).

---

[1] H. R. KRUYT and H. J. EDELMAN, *Kolloidchem. Beihefte*, 36 (1932) 350.

[2] H. R. KRUYT and H. G. BUNGENBERG DE JONG, *Kolloidchem. Beihefte*, 37 (1928) 1. Preliminary publication *Z. phys. Chem.*, 100 (1921) 250.

the agar sol bear indeed a negative charge. Further it was shown that added salts really decrease the electrophoretic velocity of agar.

Thus it seemed quite natural to correlate the sharp fall in viscosity by added salt with the decrease of the charge on the kinetic units, the higher relative viscosity of the original agar sol being caused by the electric charge of the kinetic units.

For this increasing effect of the charge on the viscosity the term "*Electroviscous effect*" was proposed.

If the conclusion drawn is correct, a similar behaviour of other negative "hydrophilic" sols as to the influence of indifferent salts on the relative viscosity was to be expected. Indeed the electroviscous effect has been found to exist in all cases investigated, viz., gum arabic [1], soluble starch [2], mucilage of semen lini [3] and of carrageen [3], sodium thymus nucleate [4] and sodium yeast nucleate [5]. Pecularities shown by some of these examples will be discussed later p. 223 § 10 and p. 227 § 11.

It may be expected that electroviscous effects are not restricted to the colloids of acidic nature, summed up in § 1c (p. 186), but must also occur in such pronounced ampholytic colloids as proteins. Here at a pH higher than the *I.E.P.* a behaviour must occur quite similar to that of the agar sol. At pH values lower than the *I.E.P.* indifferent salts must also depress the relative viscosity but as the charge is here positive, the separate curve bundles should each be characterised by the different valency of the anions of the added salts.

In the older literature depressing effects of indifferent salts on the relative viscosity of protein sols had already been published and various explanations of this phenomenon had been given. LOEB [6], experimenting with gelatin sols, could show that the valency of the cation is of primary importance for sols at a pH higher than the *I.E.P.*, and the valency of the anion for sols at a pH lower than the *I.E.P.* In the first case the gelatin sol is negatively charged, in the latter positively charged. Thus instead of the explanation LOEB gave based on the DONNAN equilibrium, these depression effects seemed to be much simpler interpretated as electroviscous effects.

KRUYT and LIER [7] using an exact technique of viscosity measurement could obtain quite normal electroviscous effects with casein sols.

The results with the negative sol (Fig. 17 a) did not differ from those with the agar sol [8]. Three bundles are obtained, each characterised by the valency of the cation. In the positive sol (Fig. 17b) only four salts were investigated. but their curves lay just in the order to be expected, the valency of the anion being of primary importance.

The only complication occurring here is that the spreading of the curves for

[1] H. R. KRUYT and H. J. C. TENDELOO, *Kolloidchem. Beihefte*, 29 (1929) 396.

[2] H. G. BUNGENBERG DE JONG, *Rec. trav. chim.* 43 (1924) 189.

[3] H. G. BUNGENBERG DE JONG and ONG SIAN GWAN, *Kolloidchem. Beihefte*, 29 (1929) 436.

[4] H. G. BUNGENBERG DE JONG and ONG SIAN GWAN, *Kolloidchem. Beihefte*, 31 (1930) 89.

[5] H. G. BUNGENBERG DE JONG and N. F. DE VRIES, *Rec. trav. chim.*, 49 (1930) 658.

[6] J. LOEB, *Proteins and the theory of colloidal behaviour*, Mc Graw-Hill Publishing Co. Ltd., London, 1922.

[7] H. R. KRUYT and H. LIER, *Kolloidchem. Beihefte*, 28 (1929) 407.

[8] The curve for NH_4Cl lies quite abnormally to the left of that for $Co(NH_3)_6Cl_3$, but the pH change (10.5 → 9.7) which this salt of a weak base (NH_4OH) brings about in the original sol accounts for this abnormal behaviour. The other salts used did not alter pH in the concentrations used. That NH_4Cl, behaves as other salts of the type $1 — 1$ in the agar sol (Fig. 16) is due to the fact that small pH changes in the neighbourhood of pH $6 — 5$, have not yet any detectable influence on the dissociation state of the acidic ionogenic groups of agar, which are ester sulphate groups.

the tri- and tetravalent anion is very small, to which peculiarity we shall return later in § 10.

The primary importance of the anion valency can be seen to apply also for positively charged ichthyocoll sols [1] from the relative positions in Fig. 18 of the curves
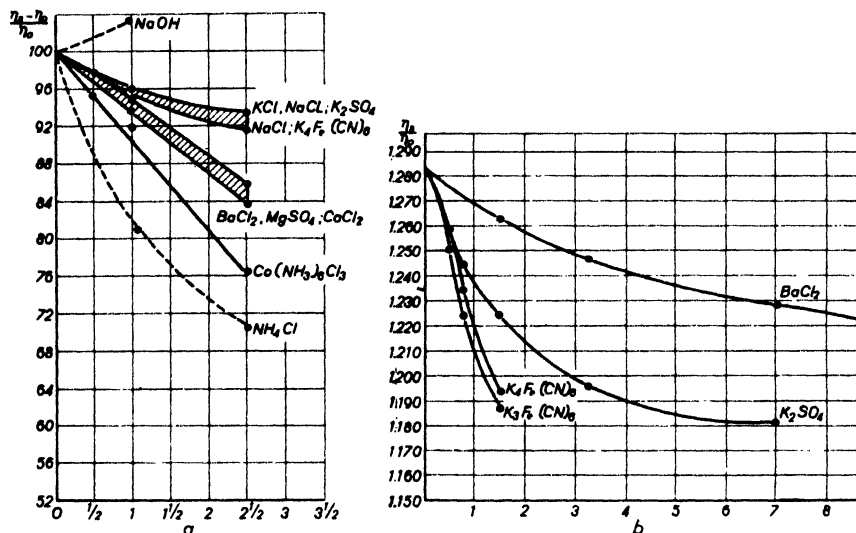
Fig. 17. Influence of salts on the relative viscosity of a negatively (a) and a positively charged (b) 1% casein sol (50° C.).

a: at pH 10.5 the curves arrange themselves in bundles characterised by the valency of the cation.
b: at pH 2.75 on the other hand the valency of the anion is determinative for the relative sequence of the curves.
In a the curve for NH_4Cl is discrepant; which results from the displacement of the pH to lower values. NaOH addition does in fact increase the relative viscosity. Flocculation occurs at the lower end of the curves for $K_3Fe(CN)_6$ and $K_4Fe(CN)_6$ (see § 10).

obtained with KCl (1—1), K_2SO_4 (1—2) and $K_3Fe(CN)_6$ (1—3). Fig. 18c shows moreover, that in these positive sols the cation influence is of no importance, the curves for $CaCl_2$ (2—1) and $Co(NH_3)_6Cl$ (3—1), nearly coinciding with the curve for KCl (1—1).

For the influence of salts on η_s/η_0 of positively charged clupein sols [2] see later in § 12.

In 1916 SMOLUCHOWSKI [3] gave an extension of EINSTEIN's formula for the case when the (undeformable) particles bear a capillary electric charge:

$$\frac{\eta_s - \eta_0}{\eta_0} = k\varphi \left[1 + \frac{1}{\eta_0 \varkappa r^2} \left(\frac{D\zeta}{2\pi} \right)^2 \right]$$

[1] H. G. BUNGENBERG DE JONG and N. F. DE VRIES, *Rec. trav. chim.*, 50 (1931) 238.
[2] H. G. BUNGENBERG DE JONG, W. A. L. DEKKER and P. VAN DER LINDE, *Rec. trav.. chim.* 54 (1935) 1.
[3] M. VON SMOLUCHOWSKI, *Kolloid-Z.*, 18 (1916) 190, c.f. also W. KRASNY-ERGEN, *Kolloid-Z.* 74 (1936) 172. F. BOOTH, *Nature*, 161 (1948) 83.

From the formula it can be seen that a positive as well as a negative charge on the particles increases the relative viscosity. For the additional term

$$\frac{1}{\eta_o \, \varkappa r^2} \left(\frac{D\zeta}{2\pi}\right)^2$$

called by SMOLUCHOWSKI the "quasiviscous effect", is intrinsically positive, the capillary electric potential ζ occurring in it in the quadratic form.
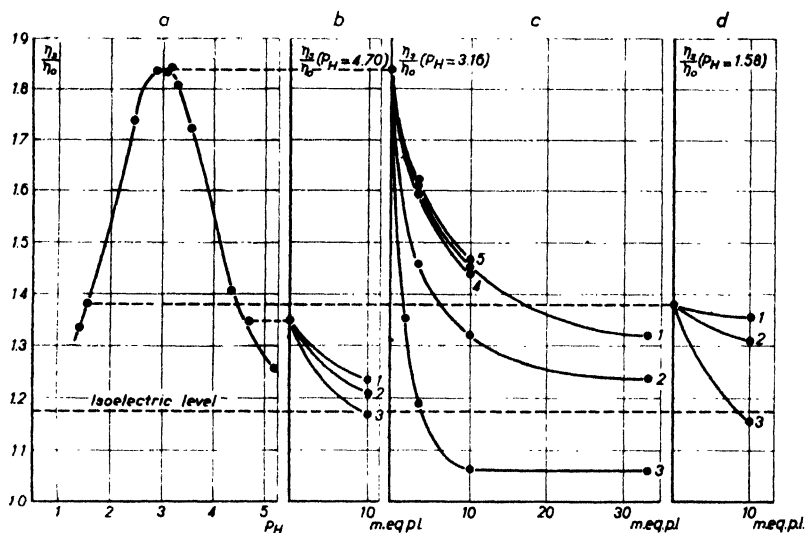
Fig. 18. Influence of salts on the relative viscosity of positive 0.42% ichthyocoll sols (37° C).
a: Part of the pH-viscosity curve at pH values lower than the I.E.P. (see Fig. 20 curve A).
b, c & d: influence of salts in which a point is chosen respectively on the rising branch, at the maximum and on the falling branch of the curve in a as working point.
1 = KCl, 2 = K$_2$SO$_4$, 3 = K$_3$Fe(CN)$_6$, 4 = CaCl$_2$, 5 = Co(NH$_3$)$_6$Cl$_3$.
The K$_3$Fe(CN)$_6$ curves fall to below the value of η_s/η_o corresponding to the I.E.P. (lowest dotted horizontal line). This is associated in c with the occurrence of highly cloudy systems, which on microscopic examination is found to be based on coacervation. (On this point see also § 10).

This theoretical deduction of SMOLUCHOWSKI seems to give at least a qualitative interpretation of the experimental results discussed above.

If we assume, that the kinetic units are undeformable, which in the small concentrations used in Fig. 16, 17 and 18 do not alter their volume, then in the above formula r the radius of the particles and φ, the total volume of the particles, expressed as a fraction of the volume of the sol, will remain constant. The only variables which may alter the relative viscosity, are then $\varkappa$, the conductivity of the dispersion medium and ζ the electrocapillary potential.

The increase of $\varkappa$ by added salts to a sol relatively free from electrolyte impurities will in itself already cause the "quasiviscous" term or — as KRUYT and BUNGENBERG DE JONG preferred to call it — the *electroviscous* term to become very small. Thus the decrease in relative viscosity by added salts within the first 100 m. eq. p. l. strictly speaking does not yet mean a discharging effect, but only that the charge can no

longer manifest itself in the relative viscosity at salt concentrations higher than 0.1 N.

The fact however, that distinct bundles of curves appear, each characterised by the valency of only one of the ions of the added salts (cations in the case of negative sols and anions in the case of positive sols), and the relative position of these bundles (the higher the valency of the ion, the lower the position of the bundle), does indicate a discharging effect of the added salts.

Summarising we may say that viscosimetry seemed to give an insight into the nature of the sol particles of hydrophilic sols: Added salts in low concentrations reveal the existence of a double layer, resembling very nearly that which is present on the particles of hydrophobic sols. At concentrations higher than 0.1 N, the electroviscous effect now being practically eliminated, large changes in relative viscosity mean changes in solvation only. These conclusions were the basis of the stability theory discussed in § 4 (see p. 198).

In the early stages of this simplistic stability theory, the removal of the electroviscous effect in soluble starch sols was studied [1]. These results seemed to draw the analogy between "hydrophobic" and "hydrophylic" sols as regards the nature of the electric charge a step further still.

In the first named sols a reversal of charge may in many cases be obtained by salts with polyvalent oppositely charged ions (see Volume I). We then meet with a zône of flocculation around the point of charge reversal („Unregelmäszige Reihen" i.e., "irregular Series"). If this reversal of charge should occur also in hydrophilic sols, no flocculation zône is to be expected according to the above named stability theory. For the hydration shell in itself would assure stability though a capillary electric charge were absent.

Assuming that such a reversal of charge would occur in such low salt concentrations, that by the smallness of $\varkappa$ electroviscous effects may

Fig. 19. Influence of salts on $(\eta_s - \eta_o)/\eta_o$ of a 1.97% sol of amylum solubile (20° C) in the region of small concentrations.
Ordinates and abscissae: as in Fig. 16. A = KCl, B = BaCl$_2$, C = Co(NH$_3$)$_6$Cl$_3$ and D = [Pt(en)$_3$] (NO$_3$)$_4$ fit completely with the behaviour in Fig. 16. E = hexol nitrate (composition, see note 1 on page 208) gives a curve with a minimum. At this minimum reversal of charge of the sol takes place from negative to positive. Flocculation does not take place with any of the salts, even not around the minimum of the hexol nitrate curve.

still manifest themselves, then the presence of such a reversal of charge will show itself in a peculiar form of the corresponding viscosity curve.

For as the "electroviscous term" (formula on page 205) contains ζ in the quadratic form, this term can be either zero (if $\zeta = 0$) or have a positive value (ζ being either negative or positive).

The curve for the relative viscosity must therefore show a minimum, which

[1] H. G. BUNGENBERG DE JONG, Rec. trav. chim., 43 (1924) 189.

corresponds to the concentration at which reversal of charge occurs. At higher concentrations the relative viscosity must thus increase, but because of the simultaneously increasing x it must soon reach a maximum and then decrease further.

This state of affairs has been shown to exist in the sol of soluble starch using hexol nitrate (a complex cobalt salt with hexavalent cation[1]).

Fig. 19 shows the results obtained. The curves for the salts with mono- di- tri and tetravalent cations, have the same relative positions and shapes as with the agar sol (Fig. 16). Hexol nitrate — in accordance with the still higher valency of the cation — initially lowers the relative viscosity still more than the complex Platinum salt, and shows further the above predicted curve form.

By electrophoresis it could indeed be shown that reversal of charge takes place. Further flocculation nowhere occurs (as also neither with the other salts). But as was to be expected, with an appropriate amount of added alcohol maximum flocculation around the minimum of the hexol nitrate curve could be obtained.

Though the underlying mechanism of the reversal of charge with proteins by changing the pH is quite different, the minimum of the well known viscosity-pH curve is comparable to the minimum discussed for soluble starch with hexol nitrate. This minimum, occurring at the *I.E.P.*, is here quite easily detectable, because of the still low values of x.

The shape of the viscosity pH curve (see Fig. 20: gelatin and ichthy-

Fig. 20. pH-viscosity curves of a 0.42% ichthyocollsol (37° C) and of a 0.67% gelatin sol (42° C). The relative viscosity is a minimum at the I.E.P. (gelatin pH 4.8; ichthyocoll pH 9) and rises on either side (fairly rapidly with gelatin, much less steeply with ichthyocoll). The viscosity falls again at sufficiently low and sufficiently high pH so that a maximum is encountered on either side of the I.E.P.

ocoll[2]) may than be interpreted as follows: Starting from the *I.E.P.* i.e., $\zeta = 0$, a pH change originates a positive or negative charge. By the smallness of x in this range electroviscous effects can manifest themselves. Thus viscosity is increased. By

---

[1] Hexol nitrate = $[Co\{(OH)_2Co(H_2N—CH_2—CH_2—NH_2)_2\}_3]$ $(NO_3)_6$.
[2] H. G. BUNGENBERG DE JONG and N. F. DE VRIES, *Rec. trav. chim.*, **50** (1931) 238.

still adding more HCl or more NaOH, $\varkappa$ becomes so great, that electroviscous effects must be gradually suppressed. The viscosity-pH curve must therefore show two maxima. See also Fig. 37 on p. 229, referring to clupein.

The different forms of the curves for gelatin and ichthyocoll in the region of the *I.E.P.*'s, which is shown in a completely analogous way in the titration curve may be explained by the supposition that the histidine content of gelatin is much larger than that of ichthyocoll. As mentioned in § 2b, the dissociation of the imidazol groups varies strongly just in this region of pH.

§ 6. REINTERPRETATION FROM THE MACROMOLECULAR POINT OF VIEW OF THE CHANGES OF $(\eta_s - \eta_0)/\eta_0$ CAUSED BY ALTERING THE COMPOSITION OF THE SOLVENT (AT CONSTANT COLLOID CONCENTRATION)

As a result of further investigations, especially those studying various properties of colloids as functions of the sol concentration BUNGENBERG DE JONG, KRUYT and LENS [1] came to an interpretation of $\dfrac{\eta_s - \eta_0}{\eta_0}$ changes, which differs considerably from the original ideas and can now be valued as a step half way to the modern interpretations. We shall not enter into detail as to the experimental facts upon which this change in ideas was based. Let it suffice to say that the sol particle was considered as a very loose structured material system of colloid substance, interpenetrated by for the greater part very loosely bound water. It was supposed that in taking up this large amount of intramicellar hydration water, still enough energy becomes available to expand the loosely structured particle very considerably until opposing elastic forces set an end to it.

All changes in $\dfrac{\eta_s - \eta_0}{\eta_0}$ were now attributed to changes in the expansion equilibrium. In "desolvation processes" the loosely built structure contracts, and the circumscribed volume of the particle decreasing, $\dfrac{\eta_s - \eta_0}{\eta_0}$ diminishes.

But also the electroviscous effect can be seen from the same point of view. For if ionogenic groups situated at the periphery of the particle dissociate, establishing a diffuse double layer, the ionised groups attached to the structure will by mutual repulsion help to expand the particle. Thus more water can be taken in till equilibrium with the opposing elastic forces is reached at a now greater volume of the expanded particle.

By adding an indifferent salt, the diffuse double layer is compressed, with the result that the repulsion between the ionised groups is diminished so that the additional expansion factor of electrical nature is greatly lessened. The swollen particle must thus expel part of the water it contains until a new equilibrium is reached at a lower total volume. The electroviscous effect and its removal by added salts are thus seen in the first place as volume changes of the expanded particles.

The picture of the sol particles as expanded very loosely built structures, has

[1] H. G. BUNGENBERG DE JONG, H. R. KRUYT, and J. LENS, *Kolloidchem. Beihefte*, 36 (1932) 429; 37 (1933) 395.

many advantages. Still it possesses a grave difficulty, which disappears however completely in introducing the randomly kinked macromolecule as its intrinsic structure.

This difficulty may be formulated in the form of the following controversy.

A. Because of the rapid decrease of hydration forces with distance a hydration layer can hardly be thicker than a few molecules of water. It is therefore very improbable that the large amount of water taken up by the dry particle is firmly bound by hydration forces. The remaining quantity must therefore rather be considered as "free" water, which is only present between the expanded structure as occlusion water.

B. The dry particle brought into water takes up a definite amount of water, shown by the definite $(\eta_s - \eta_o)/\eta_o$ values. This amount can be shifted reversibly to other definite values by changing the composition of the medium. To explain these definite amounts of water taken in there seemed no other escape than the assumption that an equilibrium is set up between expanding forces and opposed contracting forces — hydration forces versus elastic forces of the expanded particle structure. In this picture therefore we were forced to assume that all the water taken in is hydration water somehow (small amount of firmly bound water + very large amount of "diffuse hydration" water). As the residual hydration forces which bind the diffuse hydration water can only be very small, it followed that the internal structure of the kinetic unit must be very loose, enabling a great expansion before opposing elastic forces set an end to it.

In introducing the concept of the randomly kinked macromolecule as the intrinsic structure of the sol particle the above mentioned controversy no longer exists. If the dry macromolecule is brought into a medium that acts as a solvent, it will expand to the randomly kinked form described in Chapter IV, occluding a relatively large amount of solvent. The "expansion" takes place by itself, the "expanded state" being the most probable form (state of maximum entropy). Thus no work has to be done on elastic forces proper and no special forces to do this work are needed. The solvent taken in is thus not bound by solvation forces but can be described as free occluded solvent, a very small part being excluded which serves for true solvation of for instance polar groups of the macromolecule.

We have now to consider the case that the macromolecule carries ionised groups along its length. Then instead of assuming its most probable form, the macromolecule will assume as a result of the repulsive forces between the ionised groups a less probable and more voluminous form. If now pH is changed so that electrolytic dissociation of the ionogenic groups is suppressed, the volume of the macromolecule will decrease to its normal value.

A decrease of this volume will however also occur at constant pH on adding a small concentration of indifferent salts, which will diminish the mutual repulsion of the ionised groups. Here the valency of the oppositely charged ions of the added salt will be of primary importance in screening off the ionised groups.

Thus in principle the *electroviscous effect* finds a simple explanation. All other facts discussed in the preceeding section may also be explained in an analogous way.

Thus in introducing the randomly kinked macromolecule or macromolecular ion as the kinetic unit in the hydrophilic sols, (indeed all the preceding examples

of colloids do belong to this class and not to the corpuscular type, which is restricted to native proteins), volume changes of the kinetic units as postulated by BUNGENBERG DE JONG, KRUYT and LENS do certainly occur, but they need not in the ideal case to give exclusively the basis for an interpretation of the $(\eta_s - \eta_o)/\eta_o$ values experimentally found.

Indeed this volume effect enters here only as a correction factor, in the case in which the macromolecular skein is relatively dense, so that the occluded solvent acts for a part as hydrodynamically immobilized. Taking the ideal case, that the occluded solvent is not immobilized, theoretically the specific viscosity increase is given by the formula (see Chapter IV p. 108) $\dfrac{\eta_s - \eta_o}{\eta_o} \backsim N^2 . A^3 . G$ in which A is the length of a chain element, N the number of chain elements constituting the macro- molecule and G the number of macromolecules per unit of volume.

In this equation the factor NA^2 forms a measure for the density of the skein, the factor NA representing the resistance caused by the different chain elements. (See Ch. IV). If by introduction of ionised groups the density of the skein is diminished, the factor NA^2 should be replaced by a larger value. Consequently the specific viscosity is increased above the theoretical value [1].

Now in reviewing in short the changes in $(\eta_s - \eta_o)/\eta_o$ represented in the preceding figures, we must conclude that minor changes in the true hydration of polar groups situated along the macromolecule, may give an explanation of the relatively small increases in $(\eta_s - \eta_o)/\eta_o$, caused by various indifferent salts (Fig. 15 and $MgCl_2$ in Fig. 14, p. 202) in the range of higher concentration (an enforced hydration probably diminishing the flexibility of the macromolecule, cf. p. 109, 110).

The sharp fall in $(\eta_s - \eta_o)/\eta_o$ occurring with $MgSO_4$ in Fig. 14 means then, that the solvent is changed by adding this salt in a direction in which the ideal solvability, first existing, is lost. In this "salting out" medium, the macromolecule can no longer expand to the ideal statistical form, but the higher the salt concentration, the more loops of the macromolecule in the interior of the skein mutually adhere, followed soon by adhering of loops between adjacent macromolecules: the colloid substance is "salted out".

In the same way the large decrease of $(\eta_s - \eta_o)/\eta_o$ by addition of alcohol is explained (Fig. 12, p. 200). In higher alcohol concentrations the macromolecular substance is not soluble, the opalescent systems obtained with small colloid concentrations really being systems of quite other nature than originally supposed (see Chapter VIII, § 1c p. 234: "Apparent single systems of composite nature").

Further, still another factor enters here, namely the influence of alcohol in diminishing the dissociation of the ionogenic groups, this influence aiding also in the direction of interlinking loops in the interior of the macromolecule and between adjacent macromolecules (see Chapter X § 3g, p. 400).

The results obtained with tannin and simple phenols (Fig. 13, p. 201) must also be interpreted as a departure from ideal solvability. This occurs in relatively low concentrations. The renewed increase of $\dfrac{\eta_s - \eta_o}{\eta_o}$ in higher concentrations is simply

---

[1] See footnote 1 on page 185.

explained by the fact that the colloids considered are sóluble in these higher con-
centrations of tannin or resorcinol.

The macromolecular skein can in these solvents re-expand and thus $\frac{\eta_s - \eta_0}{\eta_0}$
increases to the order of magnitude, comparable with that occurring originally in
only watery medium.

The changes of $(\eta_s - \eta_0)/\eta_0$ occurring with indifferent salts (Fig. 16, 17, 18
and 19, p. 203-207) in small concentrations (the removal of the electroviscous effect)
have already been discussed above (the decreased flexibility of the chain molecule
caused by the presence of the ionised groups, is increased by the added salts).

It will not be difficult to understand also the occurrence of viscosity minima
at reversal of charge points (Fig. 20 and hexol nitrate in Fig. 19) from analogous
points of view. We shall however consider this point in more detail in § 9.

§ 7. INFLUENCE OF TEMPERATURE ON $(\eta_s - \eta_0)/\eta_0$

As discussed in Chapter IV, p. 108 theoretically $(\eta_s - \eta_0)/\eta_0$ should be
independent of temperature. A small influence of the temperature should be ascribed
to changes of the flexibility of the macromolecule. As both the direct influence of
the temperature (more intense Brownian motion, less pronounced preference for
configurations that are energetically favourable) and the indirect influence via the
hydration (higher temperature, less hydration, greater flexibility) tend to diminish
the number of monomeric groups in the chain element a small decrease of $(\eta_s - \eta_0)/\eta_0$
with increasing temperature must be expected. That in this respect a typical hydro-
philic colloid such as agar (which according to its composition as the salt of a polyose-
sulphuric acid[1], can hardly be otherwise seen as a long chain macromolecular ion)
behaves according to the expectation, is seen from the following figures [2] relating to
three agar sols. Each sol was successively measured in the order $50° \rightarrow 45° \rightarrow 40° \rightarrow$
$45° \rightarrow 50°$ C.

TABLE 2

temp. ° C.	$\frac{\eta_s}{\eta_0}$ ($^2/_7$ % sol)	$\frac{\eta_s}{\eta_0}$ ($^1/_7$ % sol)	$\frac{\eta_s}{\eta_0}$ ($^1/_{14}$ % sol)
50	2.364	1.652	1.322
45	2.393	1.662	1.325
40	2.419	1.674	1.330
45	2.393	1.663	1.326
50	2.362	1.652	1.321

From the figures may be concluded that in the short time of measurement no
irreversible changes have occurred, so that the figures may be used to calculate the
change of $(\eta_s - \eta_0)/\eta_0$ by increasing the temperature from 40° to 50° C.

We then calculate a decrease of 3.95% (2/7% sol); 3.26% (1/7% sol) and 2.49%
(1/14% sol). Obviously this percentage decrease is still a concentration function and
linear extrapolation to zero agar concentration then gives a decrease of approximately 2%.

[1] NEUBERG and OHLE, Biochem. Z., 125 (1921) 311.
[2] H. R. KRUYT and H. G. BUNGENBERG DE JONG, Kolloidchem. Beihefte, 28 (1928) 1.

§ 8. THE ELECTROVISCOUS EFFECT AS A DISTURBING FACTOR IN THE DETERMINATION OF THE LIMITING VALUE OF $(\eta_s - \eta_o)/\eta_o C$ AT COLLOID CONCENTRATION ZERO

It often has been claimed that it would be characteristic for hydrophilic colloids, that the viscosity concentration curve will deviate even at small concentrations in the positive direction from linearity. So for instance from the EINSTEIN formula the viscosity concentration curve should be linear up to values of about $\eta_s/\eta_o = 1.05$ and at higher concentrations the curve should bend upwards (see Fig. 21).

In investigating the agar sol, KRUYT and BUNGENBERG DE JONG [1], using an accurate method of viscosity measurement, found that even up to values of $\eta_s/\eta_o = 2.4$ only slight deviations from a linear
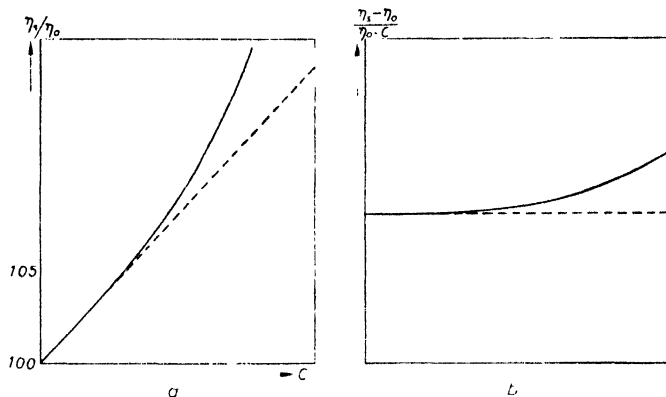
Fig. 21. Schemes to illustrate the relation to be expected between viscosity and sol concentration in the simplest case (see text).
a: the relative viscosity starts with a linear portion but at higher concentrations deviations in the positive sense occur.
b: The curve for $(\eta_s - \eta_o)/\eta_o . C$ will, proceeding toward smaller concentrations, approaches a horizontal level asymptotically.
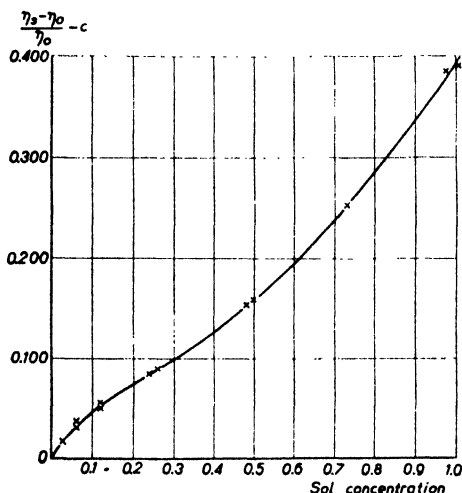
Fig. 22. Viscosity and sol concentration in agar (50° C).
Abscissae: Sol concentration expressed in terms of that of the highest concentrated sol (= 0.279% agar) as unit.
Ordinates: If the experimentally found values of the relative viscosity η_s/η_o are plotted against C, there results a relation between them which is visually near enough linear.
The curve is however S-shaped. To bring out this S-shaped character more clearly in a small scale figure the corresponding values of C (in the above mentioned sense) are each time subtracted from the values $(\eta_s - \eta_o)/\eta_o$.
Compare this figure with Fig. 21a.

[1] H. R. KRUYT and H. G. BUNGENBERG DE JONG, Kolloidchem. Beihefte, 28 (1928) 1.

function occur. These deviations had however a quite different character, the curve obtained being of the third degree, having an inflexion point at about $\eta_s/\eta_o = 1.5$

The deviations from linearity are so slight, that they will hardly be visible if the original figures of the measurements are used in a small scale figure. To enlarge them we have used in Fig 22 the values $[(\eta_s - \eta_o)/\eta_o] - C$ as ordinates. Already without introducing artifical ordinate units the deviation from expectations is plainly visible if we take $\dfrac{\eta_s - \eta_o}{\eta_o C}$ as ordinates, i.e., the increase in relative viscosity divided by the colloid concentration. This has been done in Fig. 23 (Fig. 21b is the analogous curve corresponding to Fig. 21a).

Fig. 23. The expression $(\eta_s - \eta_o)/\eta_o C$ as a function of the agar concentration.
Abscissae: agar concentration expressed in the same units as in Fig. 22.
Ordinates: The values of $(\eta_s - \eta_o)/\eta_o$ obtained from the original experimental figures, divided by the sol concentration C (in the same units as along the abscissa axis).
Compare this figure with Fig. 21b. Instead of approaching a horizontal level asymptotically as the sol concentration is lowered, the curve exhibits a minimum and seems to approach the ordinate axis asymptotically. A graphical determination of the limiting value of $(\eta_s - \eta_o)/\eta_o C$ for $C = 0$ is consequently impossible.

Thus we meet here with the experimen-

tal fact that $\dfrac{\eta_s - \eta_o}{\eta_o C}$ does not decrease to reach an end-value asymptotically on decreasing the colloid concentration (as in Fig. 21b), but increases.

The shape of the curve Fig. 23 even suggests that at still smaller colloid concentrations $(\eta_s - \eta_o)/\eta_o C$ increases further, thus leaving it quite doubtful if extra-polation to zero colloid concentration would give a well-defined value for the intrinsic viscosity.

Curve forms as in Fig. 23, or at least the left descending branch of it, have been found in all sols that were investigated: agar, gum arabic[1], soluble starch[1], sodium thymus nucleate[2], sodium yeast nucleate[3] and gelatin[4].

It has been shown, that the "abnormal" curve form is suppressed if we investigate the viscosity concentration function at a sufficient, but still relatively small constant concentration of an added salt.

[1] H. R. KRUYT and K. C. WINKLER, Kolloidchem. Beihefte, 32 (1931) 374.
[2] H. G. BUNGENBERG DE JONG and ONG SIAN GWAN, Kolloidchem. Beihefte, 31 (1930) 89.
[3] H. G. BUNGENBERG DE JONG and N. F. DE VRIES, Rec. trav. chim. 49 (1930) 658.
[4] H. G. BUNGENBERG DE JONG, H. R. KRUYT, and J. LENS, Kolloidchem. Beihefte, 36 (1932) 429.

As an example Fig. 24 gives the results for the thymus nucleate sol, the upper curve relating to sols without added salt, the lower to sols containing 5 m. eq. p. l. KCl.

This latter curve corresponding to the type of Fig. 21b is thus quite normal. The change of "abnormal" to "normal" curve forms is thus obtained in the region of salt concentrations in which the electroviscous effect is suppressed for the greater part.

The abnormal curve form for the blank can thus be explained by assuming that the electroviscous effect, being large at small colloid concentration, decreases by the mere increase of the colloid concentration. In a later investigation by KRUYT, BUNGENBERG DE JONG, and LENS[1] on the viscosity of gum arabic as a function of sol concentration, the viscosity concentration function for sols without and with added electrolytes was studied once more. It was found, that in the presence of sufficient salt a very simple relation[2] exists:

$$\log \frac{\eta_s - \eta_0}{\eta_0\,C} = k_1 C + k_2$$

the logarithm of the specific viscosity increase being a linear function of the colloid concentration (see Fig. 25). This simple relation holds for the whole concentration range studied (up to 30% gum arabic), and it applied also perfectly for the viscosity of solutions of various high polymers without electrolyte character.

Gum arabic sols without added salt gave in the range 0—10% colloid concentration a strong deviation (Fig. 25) whereas at higher colloid concentrations the above formula applies here also. Compare also Fig. 26 which refers to only small colloid concentrations (0.028—2.558%) without added salt and in the presence of small KCl concentrations.
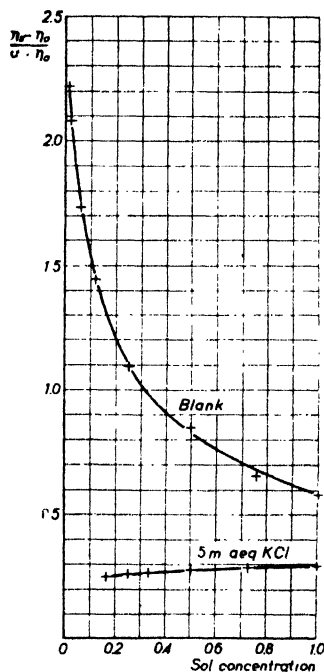
Fig. 24. $(\eta_s - \eta_0)/\eta_0\,C$ as a function of the concentration of the Na thymus nucleate sol (42° C). Abscissae: nucleate concentration expressed in that of the highest investigated concentration (= 0.164%) as unit.
Upper curve: experimental results of the sols diluted with distilled water.
Lower curve: similar experimental results for which care is taken that 5 m. eq. p. l. KCl is always present in the sols (see text.)

---

[1] H. G. BUNGENBERG DE JONG, H. R. KRUYT, and J. LENS, Kolloid Beihefte, 36 (1932) 429.
[2] This relation may be compared with the very analogous relation

$$\frac{\eta_s - \eta_0}{\eta_0\,C} = k_2 + k_1 C$$

which describes the behaviour of dilute solutions of uncharged macromolecules very well. Cf. R. H. EWART, in: Advances in Colloid Science II, 197, Interscience, New York 1946.

The logarithmic equation mentioned above however is valid up to very high concentrations (at least 30%) whereas the linear equation which has a better theoretical foundation is valid only up to concentrations of a few per cent.

In this first concentration range (up to 10% gum arabic) we have therefore once more the same abnormality as already discussed above occurring in agar (Fig. 23), thymus nucleate (Fig. 24) and other sols. It is caused by the presence of the electroviscous effect at low and the self-suppressing of it at high colloid concentrations.

Indeed as regards the influence of indifferent salt on less and more concentrated gum arabic sols, the fall in relative viscosity becomes smaller the higher the colloid concentration[1]. Compare the curve bundles in Fig. 27.
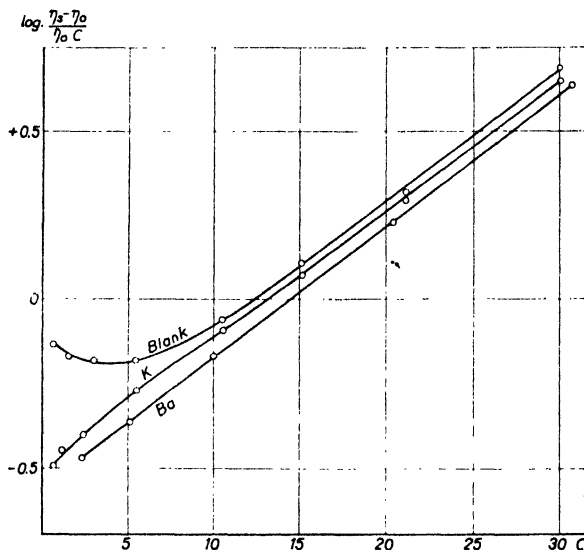
Fig. 25. log $(\eta_s - \eta_0)/\eta_0 C$ as a function of the sol concentration in gum arabic (42° C).
Abscissae: gum arabic concentration expressed in gram per 100 ml.
Upper curve: experimental results of the sols diluted with distilled water.
The curves below; ditto at constant presence of 10 m. eq. p. l KCl or BaCl$_2$.
The abnormal shape of the blank curve has given place to a simple linear course, which offers a chance of extrapolation to the sol concentration = O. The slight deviations from the linear course of the KCl curve at the smaller concentrations are possibly derived from experimental errors. The downward bend is not present in Fig. 26.

Thus it seems, that already crowding together of the sol particles into a smaller volume (i. e., increase of colloid concentrations) removes part of the large electroviscous effect existing at low colloid concentrations. This crowding together has thus in principle the same result as the addition of an indifferent salt at constant colloid concentration.

Though for these high sol concentrations the simple formulation $(\eta_s - \eta_0)/\eta_0 \infty$ $N^2 . A^3 . G$ will no longer hold, it may be supposed that among the factors determining the actual (high) viscosities, the flexibility of the macromolecule will still play a similar rôle.

H. G. BUNGENBERG DE JONG, H. R. KRUYT, and J. LENS, *Kolloidchem. Beihefte*, 37 (1933) 395.
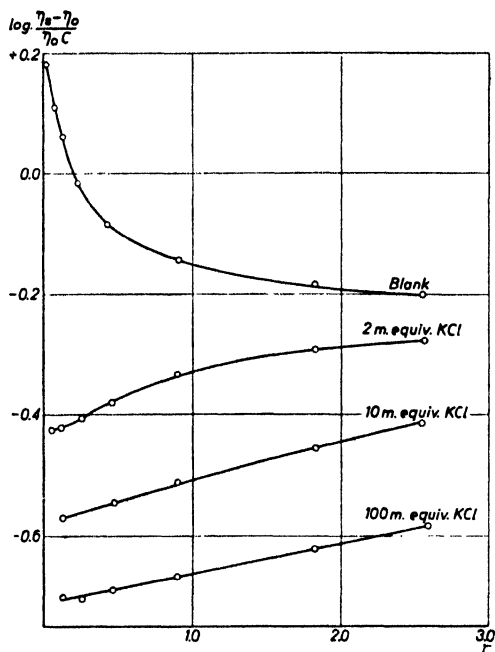
Fig. 26. log $(\eta_s - \eta_o)/\eta_o\,C$ as a function of the sol concentration with gum arabic (25° C). Abscissae: gum arabic concentration expressed in gram per 100 ml.
Top curve: without added salt. On account of the shape of the curve an extrapolation to the sol concentration = O is very uncertain.
Below: The abnormal course of the blank curve disappears on addition of sufficient KCl, as a result of which the above mentioned extrapolation does appear to be permissible.
(at 2 m. eq. p. l KCl the course of the curve is not yet quite linear).

For an explanation of the "self suppression" of the electroviscous effect we may then suppose that the macro-molecular electrolyte itself acts as indifferent electrolyte and thus diminishes the interaction between its own charged spots thus leading to a denser coil and a smaller electroviscous effect.
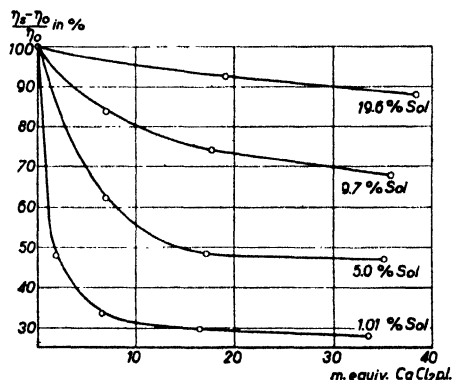
Fig. 27. Neutralisation of the electroviscous effect by $CaCl_2$ for a number of gum arabic sols of increasing concentrations.
The sol concentrations are expressed in grams per 100 ml.

§ 9. REVERSAL OF CHARGE AND ELECTROVISCOUS EFFECT

The occurrence of a viscosity minimum in the viscosity-pH curve of gelatin (see p. 208 Fig. 20) can be simply explained from the macromolecular point of view from changes in the flexibility of the macromolecule, for both a positive and a negative charge will diminish this flexibility and thus $(\eta_s - \eta_o)/\eta_o$ will then be greater than in the uncharged condition.

However we must thereby take into consideration that according to modern points of view an amphoteric protein is not really uncharged at the I. E. P., but contains an equal amount of positively and negatively ionised groups.

If ionised groups of opposite sign are thus present side by side on the macro-
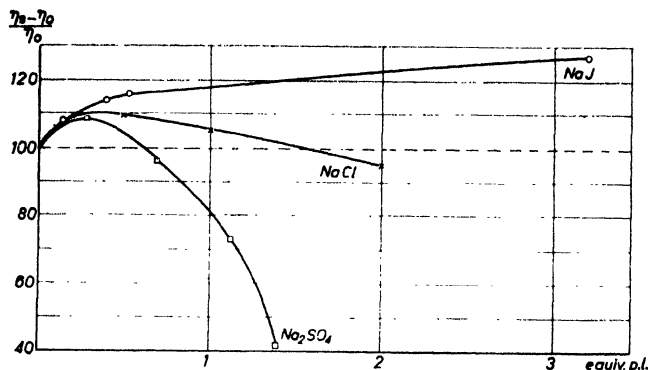
Fig. 28.   Influence of NaI, NaCl and Na_2SO_4 on $(\eta_s - \eta_0)/\eta_0$ of a 0.8% isoelectric gelatin sol (50° C). In the region of small concentrations the curves of all three salts rise above the level of the isoelectric gelatin (the value of $(\eta_s - \eta_0)/\eta_0$ corresponding to this is called 100%). At higher concentrations the curves sort themselves out into the lyotropic sequence of the anions.

molecule, they will by their mutual attraction make the skein contract below the "most probable" volume. This must have as a result that

$$(\eta_s - \eta_0)/\eta_0$$

of the isoelectric gelatin is smaller than in the really uncharged state. Added indifferent salts will by their screening effect lessen the mutual attraction of the oppositely charged groups, and thus make the skein expand again

to a volume nearer to the "most probable" one. It can thus be foreseen that added salts will increase $(\eta_s - \eta_0)/\eta_0$.

The increasing influence of indifferent salts in small concentrations on the relative viscosity of isoelectric gelatin sols found by HOLLEMAN, BUNGENBERG DE JONG, and MODDERMAN [1] can be explained in this way. Fig. 28 shows the influence of Na_2SO_4, NaCl and NaI up to relatively high concentrations.

Fig. 29 shows the same for KI, KBr, KCl and KF. In both figures so-called lyotropic influences of the anions are clearly seen, these causing the curves to bundle
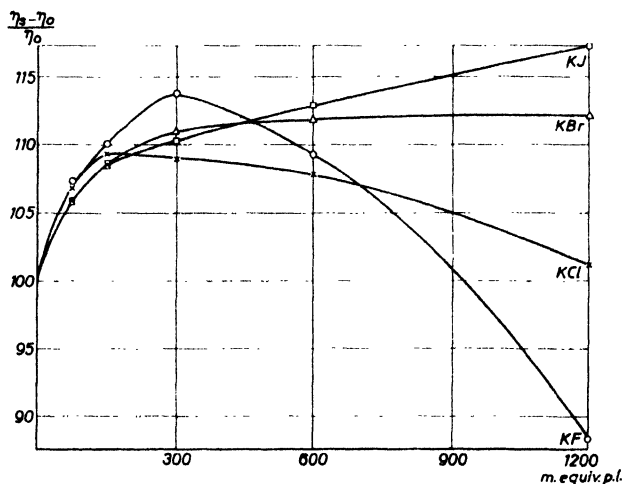
Fig. 29.   Influence of some K salts on $(\eta_s - \eta_0)/\eta_0$ of a 2.5% isoelectric gelatin sol (50° C). As in Fig. 28: at smaller concentrations rise of all the curves above the level 100%, at higher concentrations arrangement in the sequence of the lyotropic series of the monovalent anions. For the course of the KF curve, see note. 1 on page 219.

[1] L. W. J. HOLLEMAN, H. G. BUNGENBERG DE JONG, and R. S. TJADEN MODDERMAN, Kolloid-chem. Beihefte, 38 (1933) 439.
   The explanation given in this publication was not yet based on the macromolecular point of view, but already used the idea of expansion and contraction of the loosely built colloid particles.

out at higher concentrations. But we must direct our attention to the effect all salts have in common in small concentrations[1], this effect being an increase of the relative viscosity above the level for the isoelectric gelatin sol without added salts. The above given reasoning will thus explain this general effect of salts, but it foresees also the experimental fact that in this range of small concentrations not all salts will increase $(\eta_s - \eta_0)/\eta_0$ exactly in the same way.

Compare Fig. 30 which shows the influence of KCl, KNO$_3$ and KI of the type 1 — 1, further of K$_2$SO$_4$ of the type 1 — 2, of K$_3$Fe(CN)$_6$ of the type 1 — 3 and K$_4$Fe(CN)$_6$ of the type 1 — 4. By inspecting the figure, the most important factor determining the steepness of the initial increase in relative viscosity (in the concentration range 0 — 150 m. eq. p. l) is without doubt the valency of the anion.

KCl, KNO$_3$ and KI[2], salts of the type 1 — 1, together form a bundle giving the lowest increase in relative viscosity. K$_2$SO$_4$ (type 1 — 2) mounts somewhat higher (compare at 150 m. eq. p. l), whereas K$_3$Fe(CN)$_6$ (type 1 — 3) gives a still more rapidly mounting curve which is far surpassed by K$_4$Fe(CN)$_6$ (type 1—4).

Fig. 30. Influence of some salts of the types 1—1, 1—2, 1—3, 1—4 on $(\eta_s - \eta_0)/\eta_0$ of a 0.8% isoelectric gelatin sol (50° C).
All the salts produce a rise at small concentrations for which the valency of the anion is of significance. Compare the sequence of the curves at 150 m. eq.: from bottom to top there follow first the three salts of the type 1—1, then successively 1—2, 1—3 and 1—4 (see further text).

This experimentally found importance of the ion valency can be quite easily foreseen from the general explanation given above.

We must then first state, that the above reasoning will apply in its purest form only for such a salt, of which the cation will screen off the negative ionised groups of the macromolecule to exactly the same extent as its anion screens off the positively ionised groups of the macromolecule.

Such an ideal salt will not alter the isoelectric condition, and we shall assume that salts of the type 1 — 1 come near to this ideal salt[3]. With salts of the type 1 — 2, 1 — 3 and 1 — 4, the polyvalent anion will in the given order screen off the positively charged groups of the macromolecule to an increasingly greater extent than the monovalent cation screens off the negative ionised groups.

Thus also an unbalancing of the original electrical compensation characteristic of the isoelectric state, will be the result, which will increase in the order 1—2, 1—3, 1—4. Of course this unbalancing will in itself diminish the efficiency of the mutual

---

[1] The irregular course of the KF curve in Fig. 29 in the range of smaller concentrations is probably caused by an increase of the pH, KF being the salt of the weak acid HF.

[2] The very low value for KI at 75 m. eq. p. l seems suspect and may be due to an experimental error.

[3] The differences between KCl, KNO$_3$ KI in Fig. 30 show already, that besides the valency other factors of specific nature play a rôle. They will not be considered here. See Chapter IX, p. 259.

attraction of the oppositely charged ionised groups of the macromolecule and the sequence of the curves 1—4 > 1—3 > 1—2 > 1—1 in Fig. 30 may be explained by this.

Reversal of charge with hexol nitrate of negatively charged macromolecular sols — first met with in the case of amylum solubile (see p. 207 Fig. 19) — has been shown to occur very generally. In general they do not lend themselves to a viscosimetric analysis because, in contrast to amylum solubile, flocculation or coacervation phenomena in most cases accompany the reversal of charge (see p. 223 § 10).

Also lower valent cations share this property of bringing about reversal of charge, but the required salt concentrations are then increasingly higher the lower the valency of the cation (see p. 276 Chapter IX § 2).

If such a colloid and such salts are chosen, that form favourable combinations for viscosimetric investigation, — favourable here meaning that no flocculation or coacervation accompanies ·the reversal of charge — then it may be expected that, similar to the combination amylum solubile + hexol nitrate minima in the$(\eta_s - \eta_o)/\eta_o$ curves must also occur.

The viscosimetric detection of these minima for salts, which bring about reversal of charge at not very low concentrations, will however not be very easy. For at these higher salt concentrations (where activity coefficients are already low) it may be expected that the branch ascending from the minimum at the side of the higher salt concentrations will be still much less steep than for hexol nitrate in Fig. 19.

The ordinary technique of viscosimetry, which gives figures accurate to a few 0.1%, will at most be able to detect the existence of the predicted minima in the viscosity curves, but will not suffice to study them in more detail. HOLLEMAN and BUNGENBERG DE JONG[1]
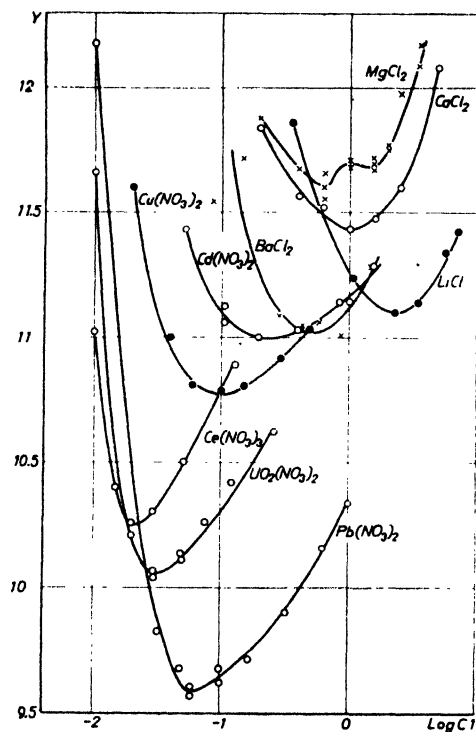
Fig. 31. Influence of a number of salts on $(\eta_s - \eta_o)/\eta_o$ of a 1% Na arabinate sol (25° C). Ordinates: Y = quantity which apart from a small correcting factor (which is however constant) corresponds to $(\eta_s - \eta_o)/\eta_o$ expressed in % of the corresponding value for the sol without added salt. Abscissae: logarithm of the salt concentration, the latter expressed in eq. per 1 (thus — 2 = 0.01 N, — 1 = 0.1 N, 0 = 1N).
With all the salts a minimum curve is obtained which lies very close to those concentrations at which reversal of charge occurs (compare with Fig. 32 and 33. See further text).

[1] L. W. J. HOLLEMAN and H. G. BUNGENBERG DE JONG, Kolloidchem. Beihefte, 46 (1937) 113.

used for that purpose a technique permitting the measurement of the time of flow with a reproducibility of a few 0.01%. With the sodium arabinate sol they could detect the existence of these minima in the relative viscosity, using one trivalent, seven divalent and one monovalent cations.

The results are given in Fig. 31, in which y represents $\dfrac{\eta_s - \eta_0}{\eta_0}$ in per cent of the blank. This quantity is given as a function of the logarithm of the salt concentration, (the use of log C is here of advantage to make the minimum character of the viscosity curves for all salts plainly visible in one and the same figure).

The reversal of charge concentrations were also determined electrophoretically on the movements of suspended SiO_2 particles in the arabinate sol, see Fig. 32.

Comparing the curve minima with the reversal of charge, a direct correlation is plainly visible, see Fig. 33. The order in which the minima follow from left to right is the same as that of the reversal of charge concentrations, but in general the latter lie somewhat (mean 0.2 units) to the right.

This investigation has thus shown that indeed minima in the relative viscosity occur as predicted for salts

Fig. 32. Reversal of charge of 1% Na arabinate with the salts of Fig. 31.
Ordinates: Electrophoretic velocities measured on suspended SiO_2 particles, expressed in arbitrarily chosen units.
Abscissae: log of the salt concentration (expressed in equiv. per l).

bringing about reversal of charge, though the systematic non-coincidence of minimum and reversal of charge experimentally found, still needs an explanation [1].
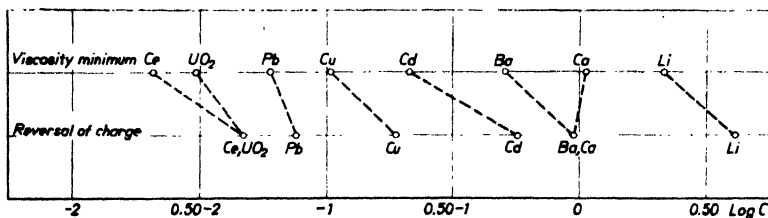
Fig. 33. Relative positions of the viscosity minima of Fig. 31 and of the reversal of charge concentrations of Fig. 32.
Abscissae: logarithms of the salt concentrations (expressed in equiv. per l).
The reversal of charge concentrations lie in general somewhat higher (average 0.2 logarithmic units).

---

[1] For instance it may be asked if the reversal of charge of the adsorbed arabinate film on SiO_2 does coincide with that of the arabinate particles free in solution, or lies systematically too high. Other possible causes are discussed in the original publication.

Of more importance are here the different levels, at which for every salt used the relative viscosity reaches its minimum value.

If at the reversal of charge point the macromolecule was really uncharged, we would expect that the minima with all salts investigated would lie on the same horizontal level in Fig. 31.

As discussed above (p. 218) for the *I.E.P.* of gelatin, we can however here also assume, that the reversal of charge state does not represent a state in which all charges on the macromolecule are absent, but one in which positive and negative charges compensate one another. Local positive charges can for instance arise from polyvalent cations fixed on the negative monovalent ionised carboxyl groups of the arabinate molecule. For monovalent cations this will not suffice, but an extra fixation on other polar groupings (e.g., OH groups) must be assumed to create local positive charges on the macromolecule.

In principle the different levels at which the minima are situated in Fig. 31 are then no longer inexplicable.

For we must assume here as above for the *I.E.P.* of gelatin that the simultaneously present positive and negative charges on the macromolecule will cause a "contraction" of the macromolecular skein, and this contraction can be of different grades, it being e.g., greater if the salt concentration at the viscosity minimum is lower (a smaller general screening effect of ions present then opposes this "contraction").

Inspection of Fig. 31 reveals that these minima are not distributed at random over the figure. On the contrary a distinct correlation between y *min* and log C (of the minimum or of the reversal of charge) exists. Low values of y *min* are found with $Ce \cdots$, $UO_2 \cdots$ and $Pb \cdots$ showing low values for log C, whereas high values of y *min* are found with $Mg \cdots$, $Ca \cdots$ and $Li \cdot$, having high values of log C.

We shall see later that reversal of charge is directly connected with fixation of cations on the ionised groups of the colloid and that as regards this fixation specific properties of the cations play a great rôle. Here no longer valency of the cation is the only important quality, but others such as volume, polarisability and polarising power of the cation come into play. See Ch. IX, § 2, p. 276.

Now in Fig. 31 such specific features also occur, the valency of the cation not being the only factor which determines the $\frac{\eta_s - \eta_0}{\eta_0}$ value at the minimum.

The divalent cations of Pb and Cu, belonging to the B subgroups of the periodic system act in this respect much more strongly than the divalent cations of Ba, Mg and Ca, belonging to the A subgroups of the periodic system, and even stronger than the trivalent cation Ce.

The specific actions of course disturb the expected correlation between height of the viscosity minimum and reversal of charge concentrations.

If we consider cations belonging to a natural family, thus the series Ba, Ca, Mg or the series Pb, Cu, Cd, then we find that indeed with a higher reversal of charge concentration there corresponds a higher $\frac{\eta_s - \eta_0}{\eta_0}$ value in the minimum.

We may conclude that the different $(\eta_s - \eta_0)/\eta_0$ values, which occur in Fig. 31 at the viscosity minima must be connected with different grades of contraction of the macromolecular skeins, caused by electrical attraction forces, set up by localised attachment of cations on the macromolecule, especially on its negatively charged ionised groups.

Such attractions may also exist between adjacent macromolecular skeins. Thus it is probable that at and around the viscosity minima associations of macromolecules exist.

With the cations studied here the arabinate sol still remains perfectly "stable", but with 4 and 6 valent cations flocculation or coacervation occurs. These higher valent cations have lower or even very low (true) reversal of charge concentrations (see p. 259 Chapter IX § 1 and 2), that means:

a. that the attachment of cations is here much stronger,

b. that the general screening effect of ions present is now much less.

Both factors will thus be favourable for producing still stronger contractions in the macromolecular skeins and much more intense mutual interactions between adjacent macromolecules. Obviously with these 4 and 6 valent cations the latter reach such a magnitude that actual transgressions of solubility occur.

§ 10. TRANSGRESSIONS OF SOLUBILITY ACCOMPANYING THE SUPPRESSION OF THE ELECTROVISCOUS EFFECT

The behaviour of soluble starch sol towards reversal of charge with hexol nitrate, without flocculation accompanying it, is rather an exceptional case (p. 208).

As a rule negative colloids of acidic nature are flocculated or coacervated by small amounts of hexol nitrate, rhodochrome chloride and $Pt(en)_3(NO_3)_4$, that are salts with 6, 5 or 4 valent complex cations (see p. 270, Table 2)

In the simplistic stability theory discussed in § 4, this "instability" occurring as a result of "discharging" at very low concentrations, does not fit in. One would have to assume, that simultaneously with the reduction of ζ a dehydration occurs, the reason why being obscure. For at these low concentrations (e.g., of a few m. eq. p. l or less) there is no reason to explain this "dehydration" as a "salting out effect", comparable for instance with the action of $(NH_4)_2SO_4$ on protein sols or of $MgSO_4$ on agar sol (see Fig. 14), which occurs as a rule at relatively high concentrations.

That these flocculations (or coacervations) with high valent cations are of a quite different nature is clearly shown in the case of gum arabic. Its sols, as POHL[1] has already shown, cannot be salted out even by the highest concentrations of salts commonly used for this purpose. Nevertheless a 1% sol is coacervated by hexol nitrate in concentrations higher than 5 m. eq. p. l, and more diluted sols at proportionally lower concentrations, the latter fact already indicating that the expression "concentration" is here of a doubtful use (for fuller information see p. 262, Chap. IX § 1 b).

Still another example may be quoted to illustrate the quite different nature of salting out and of flocculations with hexol nitrate and other salts with polyvalent ions in small concentrations: amylum solubile sols are readily salted out e.g., with certain sulphates, whereas its sols do not flocculate with the named 6, 5, or 4 valent complex cations.

Salting out must be considered as being caused by changing the solvent in such a direction, that the macromolecule becomes difficultly soluble along larger parts of

---

[1] POHL, Z. physiol. Chem., 14 (1890) 155.

its length, irrespective of the presence or absence of ionised groups. The flocculations (coacervations) with the named polyvalent ions on the other hand are primarily connected only with fixations of these ions on the ionised groups of the macromolecule and one very important factor is thereby the frequency of occurrence of the ionised groups along the macromolecule.

If this frequency is very small, as in soluble starch, even the highest valent cation will not suffice to cause flocculation or coacervation.

In gum arabic the frequency is much greater and it allows of coacervation with salts of the types 6—1, 5—1, 4—1, yet not with lower valent cations.

Other colloids have still greater frequency of occurrence of the ionised groups and here also transgression of solubility is possible with 3—1 or even with 2—1. (For fuller information on the above statements see Chapter IX in particular § 1 d, p. 269 and Chapter X in particular § 3 d, p. 392).
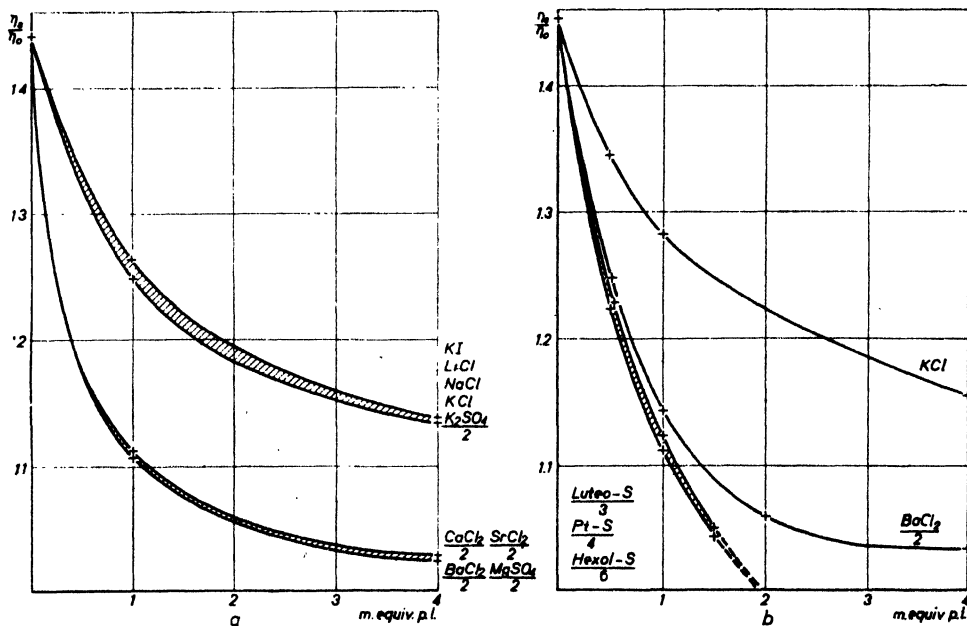
Fig. 34. Influence of salts in small concentrations on the relative viscosity of 0.1% Na thymus nucleate. (42° C).
Ordinates: relative viscosity.
Abscissae: salt concentration in m. eq. p. l.
a: The given salts divide into two groups, characterized by the valency of the cation. K_2SO_4, although of the type 1—2, acts as strongly as the four given salts of the type 1—1. In the same way $MgSO_4$ (2—2) fits with the three salts of the type 2—1.
b: The normal grouping of the curves according to the valency of the cation is restricted to 1—1, 2—1 and 3—1 but 3—1, 4—1 and 6—1 are practically equivalent and with these flocculation occurs at the bottom of the curves.
Luteo-S   abbreviation for $Co(NH_3)_6Cl_3$ (type 3—1).
Pt-S   ,,   ,, $[Pt(en)_3](NO_3)_4$ (type 4—1).
Hexol-S   ,, $[Co\{(OH)_2 Co(en)_2\}_3](NO_3)_6$ (type 6—1).
At very small concentrations $BaCl_2$ also fits closely with 3—1, 4—1 and 6—1.

An interesting feature in the course of the viscosity curves for such cases in which flocculation (or coacervation) accompanies the depression of the electroviscous effect, is the crowding together of these curves. As examples in Fig. 34 we give the influence of some indifferent salts on the relative viscosity of the sodium thymus nucleate sol[1] and in Fig. 35 the same for the carrageen sol[2].

Fig. 35. Influence of salts in very small concentrations of the relative viscosity of 1/90% carrageen sol (42° C). Ordinates, abscissae, luteo-s, Pt-s and hexol-s as in Fig. 34.
The curves start at $\eta_s/\eta_o = 2.3$ for the salt concentration zero; no measurements have however been made in the concentration range 0—0.22 m. eq. per l.
In the shaded strip A lie the curves for: LiCl, NaCl, KCNS, KCl, K_2SO_4, $K_3Fe(CN)_6$ and $K_4Fe(CN)_6$ thus salts with monovalent cations (types 1—1, 1—2, 1—3, 1—4).
In the shaded strip B lie the curves for: $Ba(CNS)_2$, $Ba(NO_3)_2$, $BaCl_2$, $Mg(NO_3)_2$, $MgCl_2$, $MgSO_4$, $CaCl_2$ and $SrCl_2$, thus salts with divalent cations (types 2—1 and 2—2).
Luteo—s (3—1), Pt—s (4—1) and Hexol—s (6—1) act practically equally strongly and give flocculation at the lower end of the curves.
The three dotted curves run quite discrepantly which is attributable to disturbances, connected with the strong tendency to hydrolysis of the salts in question.
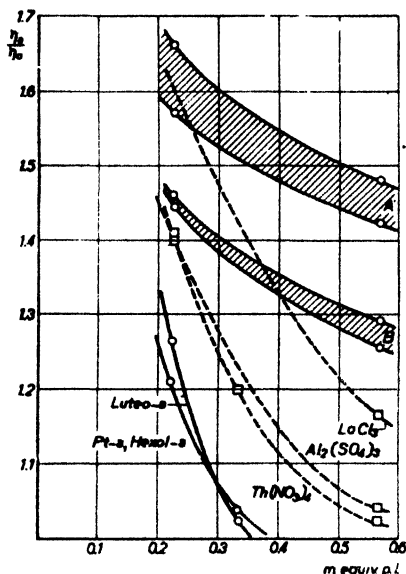
In the first named sol, we see in Fig. 34b the normal sequence 1 — 1 ... 2 — 1 ... 3 — 1 ..., but 3 — 1, 4 — 1 and 6 — 1 give approximately the same curve and all three flocculate the nucleate sol. The curves here fall to extremely low values of

$$\frac{\eta_s - \eta_o}{\eta_o}$$

In principle the same is shown by the carrageen sol, though here the bundles corresponding to the salts with monovalent or divalent cations are very broad. Here also 3 — 1 ... 4 — 1 ... and 6 — 1 form one bundle, ending in flocculation and very low values of relative viscosity.

This carrageen sol is interesting in further respects. It belongs to the extremely high viscous type, the 1/90% sol showing a relative viscosity of 2.3. The very great electroviscous effects may partly be caused by the small sol concentration (see § 8), partly be connected with the abnormal viscosity character, this sol not following POISEUILLE's law, the viscosity varying in the domain of the low shearing stresses commonly used in the OSTWALD viscometer.

For an explanation of the crowding together or even practical coincidence of the curves for polyvalent ions, we must first state that this recalls very much the

[1] H. G. BUNGENBERG DE JONG and ONG SIAN GWAN, *Kolloidchem. Beihefte*, 31 (1930) 89.
[2] H. G. BUNGENBERG DE JONG and ONG SIAN GWAN, *Kolloidchem. Beihefte*, 29 (1929) 436.

so-called *equivalent flocculation* known in lyophobic sols (see Volume I). Electro-equivalent amounts of added 6, 4 or 3 valent cations in Fig. 34 and 35 lower the relative viscosity to the same extent, and lead to transgression of the solubility of the macro-molecular colloid. These facts are much in favour of the above postulated fixation of the cations on the ionised groups of the macromolecular colloid, as the cause of flocculations and of reversal of charge phenomena.

In general flocculation (or coacervation) sets in already before the reversal of charge point is reached. We shall later see an example in which equivalent amounts of 6 and 3 valent ions are indeed fixed at the reversal of charge point (see p. 265 Fig. 4).

As it may be assumed that the fixation of these polyvalent ions is very strong, the concentrations of the cations free in the solvent in equilibrium with the fixed cations may be very small at that gross concentration in which the viscosity curves descend in Fig. 34 and 35. In this case, though the amounts of free cations may still be very different for 6, 4 and 3 valent cations, nevertheless practically every cation added to the sol may be fixed on the colloid. So that the curves may practically coincide and flocculation may occur at the same apparent concentration.

With lower valent cations, the concentrations of the ions free in the solvent, that correspond with a certain occupation of the ionised groups with cations, are so great, that they can no longer be neglected. Coincidence of the curves for salts of the type 2—1 over the whole length with the curves for 6—1, 4—1 and 3—1 will then no longer occur. It is only at very low concentrations (where fixation of cations is relatively much stronger) that the number of the cations free in the solvent may become small compared with the number fixed on the colloid. This explains the fact that in Fig. 34 at low concentrations the $BaCl_2$ curve comes very near to the bundle containing the curves for 6—1, 4—1 and 3—1.

With KCl, the conditions are still more unfavourable, so that the KCl and $BaCl_2$ curve do directly take different courses.

Reviewing the action of salts on negatively charged colloids, we may discern three cases:

A. *Simple screening effect* by the cation of the added salt, this effect giving a typical spreading of curves in the suppression of the electroviscous effect according to the valency of the cation. With increasing salt concentration sooner or later, depending on the nature of the cation and the nature of the colloid, this effect emerges into B.

B. *Fixation of cations on the ionised groups.* This effect gives crowding together of curves in the suppression of the electroviscous effect, is often accompanied by transgression of solubility, and can further lead to reversal of charge phenomena.

C. *Salting out effet*, for which the presence of ionised groups on the macromolecule is not a *conditio sine qua non* as in A and B. Relatively large concentrations are needed (see the bend in the $MgSO_4$ curve in Fig. 14 on p. 202) and the position of both cation and anion in particular of the latter, in the lyotropic series is here of primary importance.

For the action of salts on positively charged colloids the same division into three cases applies, for A and B now the valency and specific[1] properties of the anion being of primary importance. As for C (salting out) as well the action of the anion preponderates over the action of the cation, it may sometimes be difficult to discern if a flocculation or coacervation of a positive sol, e.g., with monovalent anions, is the result of the actions B or C.

Discrimination is then often possible by determining the relative intensities of flocculation with other monovalent anions.

Indications for C are then, that the order of increasing flocculation is found: CNS — I — Br — Cl — (F), the concentration needed being as a rule not very small. Indications for B, that the order is just the reverse (F) — Cl — Br — I — CNS, the concentrations needed being here sometimes relatively small (See p. 299, Chapter IX § 3 b and p. 407, Chapter X § 3 j).

§ 11. VISCOUS BEHAVIOUR OF SHORT CHAIN-MACROMOLECULAR "COLLOIDS"

Sols of long chain macromolecular colloids may be of the low viscous type if the kinetic units in their sols consist of densely built corpuscules — the long chain molecule being folded up — as is assumed to hold for many native proteins.

If the kinetic unit is not built up in that way, the macromolecular colloids belong to the high viscous type, their long chain molecules forming statistical skeins in solution.

From the formula $(\eta_s — \eta_o)/\eta_o \backsim N^2 A^3 G$ it may be seen that in a "homologous polymeric series" viscosity will increase with increasing molecular weight, (which is proportional to NA). The earliest members of such a series however must necessarily give low viscous sols. Such short chain molecules are for instance present in *Na-yeast nucleate* (which is generally formulated as a tetranucleotide, consisting of only four mononucleotides linked together by ester phosphate linkages. *Na-thymus nucleate* is constituted on similar lines (abstraction made from certain details, which are not important here) but here not only four nucleotides but a very great number are linked together by ester phosphate linkages (see formula on p. 188 § 1c).

Thus it is interesting to compare the viscous behaviour of yeast nucleate and thymus nucleate.

We have already seen in § 10, that the latter is indeed a colloid of the very high viscous type (compare the relatively high values of η_s/η_o at very small sol concentration), this being in agreement with the high polymeric nature of its molecules.

The large decrease in η_s/η_o by suppressing the electroviscous effect, is also to be expected, considering that in the backbone of the molecule — probably $(C\text{-}C\text{-}C\text{-}O\text{-}P\text{-}O)_n$ — the phosphorus atoms recurring at short distances are the centres of monovalent ionised phosphate groups. The macromolecule may then be symbolized by —⊖—⊖—⊖— etc., that is a long chain molecule in which negative charges are situated at short intervals. This of course will contribute to a markedly diminished flexibility of the macromolecule.

---

[1] The normal crowding together but now of polyvalent anions, is to be seen in Fig. 17b (p. 205), with the positive casein sol for the curves obtained with $K_3Fe(CN)_6$ and $K_4Fe(CN)_6$. They also end in flocculation.

It may even occur, that specific properties are of so great importance that an inversion of the normal sequence of the curves in the suppression of the electroviscous effect is the result. See Fig. 38 (p. 230) in which $K_3CH(SO_3)_2$ (type 1—3) decreases η_s/η_o more than $K_4Fe(CN)_6$ (type 1—4).

Thus a large electroviscous effect may be expected and consequently a large decrease in relative viscosity on adding indifferent salts (see p. 224 Fig. 34).

Now turning to Na-yeast nucleate, we meet with a "colloid" which only slightly increases the viscosity of the solvent. Compare Fig. 36, where we see that the 0.83% sol has only a relative viscosity of 1.052 (The 0.1% thymus nucleate however gives $\eta_s/\eta_0 = 1.45$).

For an accurate measurement of the influence of salts on the viscosity of yeast-nucleate it was necessary to use a technique which gave a reproducibility of a few 0.01% in the time measurements.

The results[1], see Fig. 36, gave a normal sequence of the curves for KCl, BaCl$_2$ and Co(NH$_3$)$_6$Cl$_3$ and we may interpret this decrease of the relative viscosity as a suppression of the electroviscous effect.

The yeast nucleate molecule is however relatively so small that its four mononucleotide residues possibly do not suffice to form a chain element. Thus a skein proper will not be present in solution. Nevertheless in principle the decrease
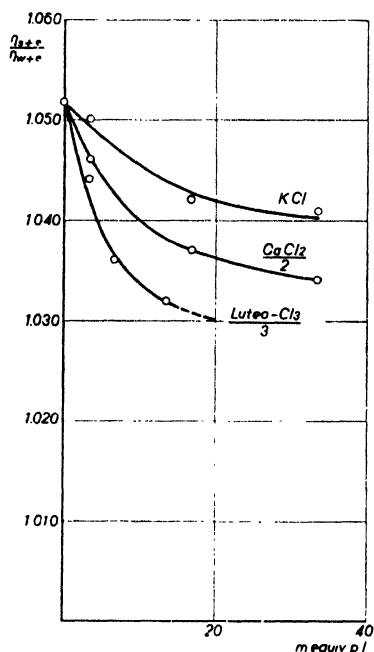
Fig. 36. Neutralisation of the electroviscous effect for a 0.83% sol of Na yeast nucleate (42° C). Beyond 13.3 m. eq. per l. Luteo—s (= Co(NH$_3$)$_6$Cl$_3$) flocculation occurs. This figure should be compared with Fig. 34 from which it appears that, notwithstanding the fact that yeast nucleate belongs to the low viscous, thymus nucleate to the very high viscous type of macromolecular colloids, the character of the bundle of curves is the same and also Co(NH$_3$)$_6$Cl$_3$ causes flocculation with both (and this occurs with both with CaCl$_2$ also at higher concentrations, see p. 270).

in viscosity by added salt may be explained on similar lines from the increased flexibility of the (short) chain molecule. Without added salt the latter will by the mutual repulsion of the ionised phosphate groups assume a more or less stretched form. If this "electrical stiffening" is removed by added salts, which screen off the ionised phosphate groups, the chain molecule may assume a more bent form.

It is interesting that yeast nucleate also gives flocculation with Co(NH$_3$)$_6$Cl$_3$. Later (see p. 395) we shall see that as regards transgression of solubility by adding different cations the low polymeric yeast nucleate and the high polymeric thymus nucleate behave very similarly.

§ 12. LOW VISCOUS BEHAVIOUR AS A RESULT OF EXISTING "COMPLEX RELATIONS"

Investigations of the viscous behaviour of clupein sulphate sols have shown that these also give relatively low viscous sols, so that here also the measurement had to be

[1] H. G. BUNGENBERG DE JONG and N. F. DE VRIES, Rec. trav. chim., 49 (1930) 658.

performed using a technique allowing reproducibility of time measurements to a few 0.01%.[1]

The viscosity-pH curve (Fig. 37) shows in principle the expected form, if one takes into consideration that the *I.E.P.* is situated at a high pH value, the only point which calls for explanation being the very low order of the η_s/η_o values which are obtained.

Fig. 37. pH-viscosity curve of a 0.292% clupein sulphate sol (25° C). (See text). The ordinate gives the quotient of the times of flow in an OSTWALD viscometer for the sol (t_{s+e}) and a liquid of the same electrolyte composition but containing no clupein (t_{w+e}). As the addition of the clupein has only a slight influence on the density, the ordinate gives practically the relative viscosity η_s/η_o. The shape of the curve for this colloid of the low viscous type is in principle the same as those of proteins of the high viscous type (comp. Fig. 20). Minimum at the I.E.P. and fall of the curves at very high or very low pH values. As a consequence of the extreme position of the I.E.P. at already fairly high pH values, the maximum in the curve to the right of the I.E.P. is here very narrow, the maximum left of the I.E.P. on the other hand very broad.

The very flat course of the curve between pH — 10 and pH = 4 and the extra bump at pH 3 can be directly explained from the amino acid composition of the clupein. Between pH — 10 and pH — 4 all guanidino groups are positively and the one carboxyl group negatively charged. The charge and thus the viscosity is constant in this range. In the neighbourhood of pH 3 the dissociation of the carboxyl group is suppressed so that there the charge increases by one unit and the viscosity rises somewhat before it decreases again at lower values of the pH, due to high concentration of the ions present.

Turbidity (which is due to coacervation) occurs at the I.E.P.
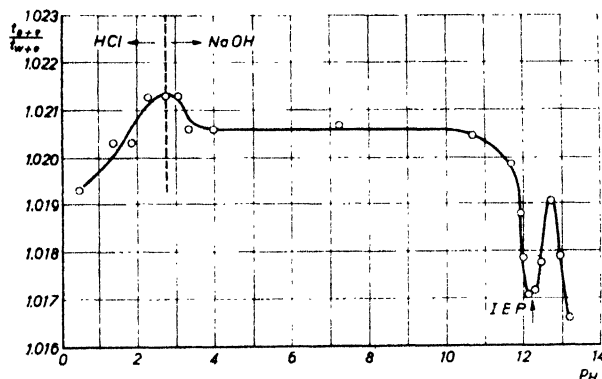
There are sufficient indications that this low viscous character is not connected with its kinetic units being densely built corpuscles as in native proteins. There are for instance the absence of denaturation phenomena (by heat or otherwise) and the great ease with which coacervates of a distinctly liquid character can be obtained. Compare Ch. VIII § 3, p. 247 and Ch. X § 3j, p. 406.

The low viscous character could further be connected (as above in the case of Na-yeast nucleate) with the macromolecules being very short, consisting only of a few monomeric residues. This once more does not strictly apply for clupein, the latter though a protein of relative low molecular weight (a few thousands) still consists of more than 30 amino acid residues linked together by peptide groups.

We may expect that this comparatively small number of monomeric residues will not give very viscous solutions, but we hardly think it possible that the very low η_s/η_o value found experimentally can be explained wholly in this way.

Considering that $\frac{2}{3}$ of the amino acids consists of arginine, we might expect reasonable depressions of the relative viscosity on adding salts.

---

[1] H. G. BUNGENBERG DE JONG, W. A. L. DEKKER, and P. VAN DER LINDE, *Rec. trav. chim.* 54 (1935) 1.

For assuming a more or less regular distribution of the arginine side-chains along the macromolecule the latter may be symbolised by

$$\overset{\oplus}{\underset{\oplus}{\mid}} \quad \overset{R}{\underset{\oplus}{\mid}} \quad \overset{\oplus}{\underset{R}{\mid}} \quad \overset{\oplus}{\underset{\oplus}{\mid}} \quad \overset{R}{\underset{}{\mid}} \quad \text{etc.}$$
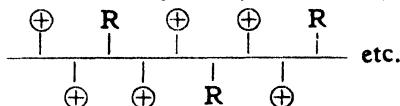
Then considerable electrical stiffening of the macromolecule may be expected (as was discussed in § 11, for thymus nucleate).

Fig. 38. Influence of small salt concentrations on the relative viscosity of a 0.32% clupein sulphate sol (pH 2.76, 25° C).
Some deviations, which are discussed in the text, here occur from a normal course of the curves in the suppression of the electroviscous effect of a positively charged sol.

Addition of salts will then increase the flexibility of the macromolecule, now decreasing considerably the number of monomeric residues constituting a chain element.

Investigation of the influence of salts on the relative viscosity (see Fig. 38) shows indeed a fall of η_s/η_o with some salts, but this depression is certainly not very great.

The sequence $KCl — K_2SO_4$ is normal, which salts did not coacervate the clupein sol at the concentrations used.

The decrease in η_s/η_o is greater with $K_3CH(SO_3)_3$ (1—3) and $K_4Fe(CN)_6$ (1—4) but the curves end here in coacervation. As discussed in § 10 such transgressions of solubility indicate "fixation" of ions on the ionised groups of the macromolecule. In such fixations the valency of the ions is no longer the only important factor, but specific factors of the ions concerned and of the nature of the ionised groups also enter. This gives a first explanation of the reversed order of the curves 1—3 and 1—4 in Fig. 38.

It is further important that K_2SO_4 at higher, but still relatively low, concentrations than used in Fig. 38 also coacervates the clupein sulphate sol, showing that such intimate relations between anions and positively charged ionised groups of clupein are already possible in the case of sulphate ions.

This gives us a clue for explaining the low viscous character of the clupein-sulphate sol. What we said above about the electrical stiffening of the macromolecule does not apply; this could only be the case if the sulphate ions were to be found at relatively great distances from the positively charged arginine side-chains. On the contrary we have to assume that already in the original sol an appreciable fraction of the sulphate anions take up positions very close to the positively charged ionised groups of the clupein (what in the previous paragraphs was called "fixation"), which necessarily will contribute to rolling up of the macromolecule into a relatively dense skein.

In this direction the divalence of the sulphate ions helps, as one SO_4 ion can then connect two monovalent positive ionised groups of the clupein molecule.
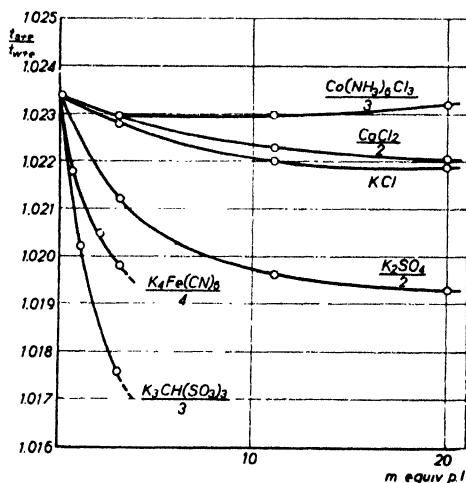
We might thus explain the low viscous character by the existence of so-called "complex relations" (see Chapter X: Complex Colloid Systems) between the positive ionised groups of the macromolecule and the dissociated sulphate ions [1].

In accordance with it is the fact, that clupein sulphate is only slightly soluble in water, addition of more clupein sulphate to its saturated solution giving two-phase systems, a liquid rich in clupein sulphate (coacervate) and a liquid poor in it (see p. 406 Chapter X § 3j).

A further detail in Fig. 38 also in accord with the supposition that complex relations are present in the original sol is the non-coincidence of the curves for 1—1, 2—1, 3—1, (compare Fig. 18c on p. 206, in which the curves for KCl, $CaCl_2$ and $Co(NH_3)_6Cl_3$ are very close together, thus showing the normal behaviour).

Considering in Fig. 38 the relative position of these curves and those for 1—2 and 1—3, we easily recognize the sequence of the so-called "continuous valency rule".

$$1—3 \ldots 1—2 \ldots 1—1 \ldots 2—1 \ldots 3—1$$

which, as will be discussed in Chapter X (p. 352, 388, 408, 415), is characteristic of systems in which complex relations are present. In the nomenclature of that chapter the clupein sulphate sol should be named a "Complex sol" (p. 336, 337).

---

[1] In most other cases complex relations between colloids and oppositely charged ions arise only upon addition of electrolyte to the sol. In the case of the clupein sulphate, the complex relation is strong enough to be manifest in the pure sol.

VIII. CRYSTALLISATION — COACERVATION — FLOCCULATION

H. G. BUNGENBERG DE JONG

Department of Medical Chemistry, University of Leyden

§ 1. LOW AND HIGH DISPERSED STATES OF THE COLLOID-RICH PHASE WHICH SEPARATES OUT

a. Crystallisation and Coacervation

Fig. 1. Crystals of pepsin. 90   lin (according to HERRIOTT).

If one starts from a sol, that is the solution of the colloid in an appropriate solvent, then according to the nature of the colloid, various changes (temperature, pH, addition of a substance) can bring about a reduction of the solubility as a result of which a larger part of the colloid separates out in a new phase.

The original one-phase system — the sol — thus divides into two phases, one of which is rich in colloid, the other poor.

The separated colloid-rich phase can either appear in a low dispersed state or in higher dispersed states. In the first case macroscopic or microscopic investigation allows one to distinguish between *crystallisation* [1], when obviously crystalline individuals (see Fig. 1) are formed and *coacervation*, when amorphous liquid drops are formed (see p. 235,

[1] With regard to „paracrystals" see p. 243.

Fig. 4), which latter coalesce more or less readily (see Fig. 2 and on p. 444 Fig. 11) and may under favourable conditions even coalesce in a relatively short time (order of hours) into one clear homogeneous colloid-rich liquid layer (coacervate layer, see Fig. 3).
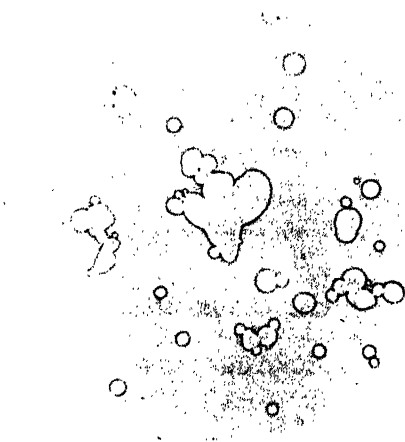
Fig. 2. Coacervate from serum albumin (positive) + gum arabic (negative). 120 × lin. The coacervate drops have partially flowed out over the surface of the micro slide and so coalesced with each other.
The coacervate belongs to the type of complex coacervates which are further discussed in the Chapters X, § 2 (p. 338) and XI (p. 433).
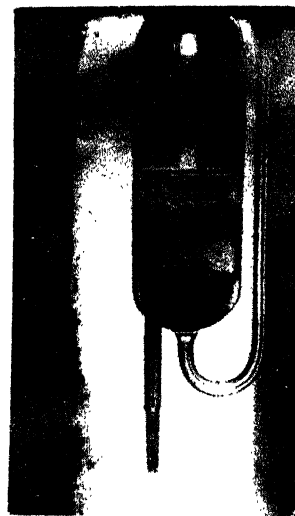
Fig. 3. Example of complete separation into two macro-layers after coacervation. ³ ₐ lin. (Isoelectric gelatin sol + isopropyl alcohol at 42)
A thermometer inserted in the thermostat behind the glass vessel can be seen through both layers. The displacement which the thermometer shows in going from the colloid poor layer (upper) to the coacervate layer (lower, results from a difference in refractive index of the two layers

b. Flocculation and formation of apparent single colloid systems

In the cases which are called *flocculation*, the separated colloid-rich phase is present in a *higher dispersed state* and here it frequently costs a great deal of trouble or it is no longer possible by microscopic investigation of the "floccules" to determine to which category the separated phase belongs.

The case of flocculation — in which at least the mass cohering to form floccules can be clearly distinguished macroscopically or microscopically from the colloid-poor phase — graduates via all kinds of transitions into the extreme case in which this is no longer possible.

If, for example, the separated colloid-rich phase remains for some reason or other in a very highly dispersed state of division, then systems can be produced which are macroscopically and microscopically homogeneous. Two types of these systems are known, of which the one has throughout the nature of a liquid, the other that of a solid body. They can be called *"apparent single colloid systems"*. (compare p. 7 Ch. I § 3 d). This nomenclature indicates that it is appropriate for various reasons to treat them as two-phase systems. In this respect they differ from the true *"single colloid systems"* (the original sols, the colloid crystals, the coacervates), in which the one-phase conception is the more appropriate.

The difference between the two kinds of colloid systems resides in the equilibrium character of the true "single colloid systems" and the non-equilibrium character of the "apparent single colloid systems". Considered from the two-phase standpoint, the mutual surface of contact between the colloid-rich and the colloid-poor phase is very great in these latter systems, so that they attempt to reduce this surface of contact. They will therefore change their properties with time while the true single systems do not change with time.

c. The liquid type of the apparent single colloid systems and their relation to flocculation

As an example let us compare the properties of a sufficiently dilute agar sol (above 40°, to avoid gelation phenomena) with those of the liquid system obtained by adding an excess of alcohol to this sol. The former is completely clear and shows no particles in the ultramicroscope. Addition of KCl does not result in flocculation either in small or in large concentrations.

The system mixed with alcohol is also clear but definitely opalescent. Macroscopically and microscopically it is homogeneous but with the ultramicroscope particles are seen to be present. The behaviour with respect to salts is quite different; even a small KCl concentration produces flocculation. The opalescent system containing alcohol therefore behaves as a *"lyophobic"* sol.

Electrophoretic measurements indicate that the particles are negatively charged and that this charge (ζ-potential) is decreased by added salts and indeed at the same concentrations (in m.equiv. per l) to a greater extent the higher the valency of the cation of the added salt. The same influence of valency also makes its appearance in the concentrations which are required just to flocculate the opalescent sol completely in a given interval of time.

The particles in these systems can fruitfully be regarded as merely fragments of a highly dispersed colloid-rich phase. That is to say the "hydrophobic sols" obtained behave quite differently from the original sol from which we started. This latter, being a true solution, represents thermodynamically an equilibrium state; the "hydrophobic sol" on the other hand does not represent a true equilibrium. It merely has a somewhat lengthy life on account of the capillary electric charge on the boundary surface between the phases but will, even without addition of salt, inevitably be ruined as a system by spontaneous flocculation provided one waits long enough. The first symptoms in that direction, namely a spontaneous increase of the opalescence, are clearly observable even after a short time.

Naturally this spontaneous flocculation will be accelerated by all factors which (a) increase the probability of encounter of the particles and (b) produce a decrease in the capillary electric charge (which has a relatively stabilising action).

If one chooses the concentration of the agar sol higher step by step then addition of alcohol first gives more opalescent systems, then turbid, slowly flocculating systems and finally direct flocculation. In this case factor (a) is increased. Nevertheless it is possible that factor (b) is simultaneously also in action through the accumulation of contaminating electrolyte which happens on successively choosing a higher agar concentration.

From the above it follows that the conditions for the establishment of the liquid type of apparent single colloid systems in general consist of (a) a certain charge [1] of the colloid, (b) a low sol concentration, (c) absence of electrolytic impurities and (d) the added substance which is used to reduce the solubility of the colloid must be a non-electrolyte [2].

Since, with the usual methods of reducing the solubility of the macromolecular colloid, the above mentioned conditions are as a rule not fulfilled simultaneously, the primarily produced "lyophobic sol" has practically no stability. *Flocculation* then takes place immediately afterwards, in which the "particles" of the lyophobic sol unite with each other with the formation of loosely constructed "*floccules*". The greater part of the colloid-poor phase is thus readily obtained by spontaneous sedimentation of the floccules or by moderate centrifuging. The floccules themselves are still always highly dispersed systems of the colloid-rich phase, consisting of cohering "primary particles" of the "lyophobic sol".

On account of this highly dispersed state the floccules likewise again do not represent a thermodynamic equilibrium. They will attempt to reduce the still very large boundary surface between colloid-poor and colloid-rich phase by enlarging the mutual contact spots of the "primary particles". Floccules have therefore in general the tendency to contract. Secondary processes can now also occur. When the colloid-rich phase has a coacervate nature,

Fig. 4. Coacervation of gum arabic sol with trypaflavin solution. 500 × lin.
(in the zone of contact of two drops lying together on a prepared [3] micro slide, brought into contact by applying a cover-glass).

that is to say, is liquid — although possibly very viscous — then the contact spots between adjoining very small coacervate drops can transform into coalescence spots

[1] It will thus be unfavourable if the colloid has the nature of an ampholyte and the pH is just equal to that of the isoelectric point.

[2] For other examples and properties of such "lyophobic" sols and the nature of the colloid-rich phase see H. G. BUNGENBERG DE JONG and P. VAN DER LINDE, *Proc. Koninkl. Nederland. Akad. Wetenschap.*, Amsterdam, 38 (1935) 419 (glycogen) and H. G. BUNGENBERG DE JONG, *Proc. Koninkl. Nederland. Akad. Wetenschap.*, Amsterdam, 38 (1935) 426 (starch). By heating of these lyophobic starch sols systems are produced containing very thin hexagonal plates visible in the ultra microscope.

[3] The micro slide was covered with a dried film of soluble starch, see p. 435.

and this process will only come to an end when all the floccules have "coalesced" to form a homogeneous coacervate layer. It is only when this has occurred that the thermodynamic equilibrium will be reached. Fig. 29 b and c in the Chapter X (p. 373) gives an example of such a gradual coalescence.

If the coacervate shows relatively little viscosity then an actual formation of floccules no longer takes place, since the coalescence of the cohering drops occurs fairly quickly. At each stage of the coarsening microscopic investigation then only shows detached coacervate drops of increasing size. (See Fig. 4).

If however the coacervate is very viscous and if the conditions for the maintenance of the "lyophobic sol" are not fulfilled the system flocculates and gradually a coarse-grained or lumpy precipitate is formed. An indication of the coacervate nature of a flocculation of this kind is frequently given by the tendency of the floccules to stick to the glass wall.

Finally when the coacervate is extremely viscous and more or less assumes the nature of a solid body (comparable with the glassy state) no sign of the above mentioned increases in "particle" size can be observed.

It is however also conceivable that the flocculation produced consists of cohering ultramicroscopic colloid crystals. In this case the contact surface between colloid-rich and colloid-poor phase could decrease by recrystallisation and thermodynamic equilibrium would be reached when all the very small crystalline individuals had made way for a large crystalline individual.

d. An example of the solid type of the apparent single colloid systems

The second type of "apparent single colloid systems" has the character of a solid body. It is produced from the original sol state when the separated colloid-rich phase forms one coherent mass throughout the whole volume of the original sol, as a result of which the latter sets to a *gel*. That one will frequently be :able to make use with advantage in this case also of the two-phase concept, is to be seen in the obvious opalescence which betrays the presence of relatively large regions which differ in properties from their immediate neighbourhood. The (spontaneous) syneresis also, which proves that no equilibrium prevails, can fruitfully be described as the extrusion of a colloid-poor phase from a colloid-rich phase. A good example is furnished by the agar sol on cooling.

Again the behaviour of less concentrated agar sols (0.1%) is in agreement with this. These sols no longer set on cooling to a solid jelly but flocculate, as can be established with the ultramicroscope [1]. This spontaneous flocculation is already a relatively slow process with 0.04% sols but it can be accelerated by the addition of a little indifferent salt [2].

In still more dilute agar sols (0.02%) flocculation is no longer observed even after one day. We have here again obtained a liquid "apparent single colloid system" with the behaviour of a "lyophobic sol", because if a small concentration of salt is present in this 0.02% sol then it flocculates right enough on cooling.

---

[1] BACHMAN, Z. anorg. Chem. 73 (1912) 125.
[2] H. G. BUNGENBERG DE JONG, Rec. trav. chim., 47 (1928) 797.

In agar sols of still lower concentrations these phenomena remain the same, a small concentration of salt gives flocculation. It is only at very low agar concentrations that added salts no longer have any influence. In this case cooling obviously no longer leads to the solubility limit being exceeded.

These observations lead one to describe the solid agar gel as a highly dispersed flocculation aggregate. It is not however appropriate to apply the two-phase conceptions to all gels. The degree of dispersion of the colloid-rich phases can assume such high values in other gels that it becomes inappropriate to continue to apply the concept phase to them. Thus there are also gels which one can better describe as one-phase systems (see already p. 9 Ch. I § 3 e and further the chapter on gels, p. 483 Ch. XII).

e. Recapitulation and further remarks

The points discussed above are collected in the following summary.

	Name of the process	State of dispersion of the colloid-rich phase	System obtained
A	Crystallisation	low dispersion	Colloid crystals + dilute sol
B	Coacervation	,, ,,	Coacervate + dilute sol
C	Flocculation	higher dispersion	Floccules + dilute sol
D	Formation of "lyo-phobic sols"	high dispersion	Liquid apparent single colloid system
E	Gelation	high dispersion	Solid apparent single colloid system (Gels which are still preferably to be treated as two-phase systems)
F	Gelation	very high dispersion in which the concept phase becomes inappropriate	Gels which can better be described as one-phase systems

Of these six categories E and F are discussed in detail in Chapter XII. "Gels" and of the remaining four stress will be laid on A and B for the following reasons.

A complete treatment of C and D would in fact make it necessary to study them from a combination of the diverse points of view which form the basis of Volume I and Volume II of this book.

From the point of view of Volume I we are not so much concerned with the processes which lead to the formation of a high dispersed phase. The "lyophobic sols" appeared in Volume I as given and we only sought the reasons for their "relative stability" and studied the cause and kinetics of their slow or rapid flocculation. This method applied to C and D would only further elucidate the mutual relations of C and D, C being

the flocculation product of D. A behaviour of the systems D differing in principle from those of the "lyophobic sols" described in Volume I is thus not to be expected and since furthermore systems of the type D do not occur frequently (see p. 235) this mode of study would furnish but meagre results.

From the point of view of Volume II the systems C and D do not interest us in the first place because of the high dispersed state of the colloid-rich phase but just because of the basic change of the solubility of the macromolecular substance. This change is however common to all four categories A, B, C and D.

Since now the problems, which are connected with this changed solubility, (properties and internal structure of the colloid-rich phase, its composition as a function of the composition of the colloid-poor phase) are amenable to experimental study in A and B, but only with difficulty or not at all in C and D on account of the high dispersed state, the study of crystallisation and coacervation is precisely the right way to obtain further information about the mechanism which leads to reduced solubility in macromolecular colloids.

Flocculation of the "lyophobic sols" in Volume I is usually not, or incompletely reversible (peptisation); it is in principle [1] completely reversible in the case of the macromolecular sols. If for example one has caused an agar sol (at 50°)[2], if necessary after the addition of a little salt, to flocculate by adding an excess of alcohol, then these floccules again dissolve completely on the addition of sufficient water. In the flocculation one passes successively through the changes: macromolecular solution → system D → system C.

The dissolving of the floccules on the addition of an excess of water is in this case not based on a "peptisation" (that is to say C → D) but on the colloid-rich phase going into solution.

The partial process D → C is on the other hand just as irreversible as the flocculation of the lyophobic sols in Volume I, since in this process the solubility of the macromolecular substance is not changed.

Guiding principles of a very general character regarding the factors which govern the solubility and separation into two liquid phases in macromolecular sols have already been treated theoretically in Chapter III.

Their rôle in the solubility of macromolecular colloids of non-electrolytic nature — and those are usually colloids which are not soluble in aqueous media — have already been discussed in Chapter VI.

From this it is obvious that, when an organo-sol of such a colloid is mixed with a liquid, which does not possess the character of a solvent, separation of a colloid-rich phase can take place. The latter has as a rule macroscopically the character of a flocculation. Recently cases have also become known in which macroscopic coacervation takes place [3]. In the following sections we shall restrict ourselves to crystallisation and coacervation of macromolecular bio-colloids soluble in water.

---

[1] Excluding cases in which an irreversible intramolecular reduction of solubility occurs simultaneously with or following on the flocculation. (See p. 339, Chapter X § 2c).

[2] The experiments must be made above 40° since gelation phenomena begin to appear below this temperature.

[3] A. DOBRY, *J. chim. phys.* 35 (1938), 387; 36 (1939) 102, studied in particular the coacervation of acetyl cellulose dissolved in chloroform with ethyl alcohol.

§ 2. THE UNIQUE POSITION OF THE CORPUSCULAR PROTEINS: CRYSTALLISATION AND DENATURATION

"Globular" or better [1] "Corpuscular" proteins, as already discussed earlier (Ch. VII § 1 b p. 185), occupy a position apart among the macromolecular colloids, because in them the kinetic unit consists of a tightly built corpuscle, or of an association or binding of numerous such corpuscles (submolecules), with a very specific structure. One can assume that in these corpuscular proteins the macromolecules are folded in a very specific manner into dense structural units.

This characteristic form of the kinetic units, which is so different from that of the "linear" proteins or of other macromolecular colloids explains some striking peculiarities in their behaviour.

In the first place the nature of the colloid-rich phase which separates when the conditions are such that it is obtained in the low dispersed form. Then corpuscular proteins usually [2] give colloid crystals [3], the other macromolecular colloids however giving coacervates. Secondly the frequent occurrence of irreversible changes of solubility in corpuscular proteins, which as a rule are absent in the other macromolecular colloids (or if present are based on quite different causes [4]).

The first of these differences can be sought in the ease with which these tightly built kinetic units of the corpuscular proteins can be fitted into a three-dimensional lattice on account of their well-defined three-dimensional dimensions while it seems very improbable that this would occur spontaneously when the kinetic units consist of randomly kinked [5] macromolecules.

One can hardly think of these kinetic units of the corpuscular proteins being stable otherwise than when lateral bonds (primary or secondary valencies) between groups of contiguous folds of the macromolecule or between subunits counteract to a sufficient extent the natural tendency of the long chain molecule to assume the most probable shape, that of the randomly kinked macromolecule.

Such intramolecular bonds, as long as they continue to exist, will in general hardly permit of volume changes or of changes in the relative position of the folds.

Thus, in contrast to the randomly kinked macromolecules, the solubility will not depend on all the groups in the macromolecule but only on those which lie on the surface of the tightly built kinetic unit.

---

[1] W. T. ASTBURY. *Kolloid-Z.*, 83 (1938), 130, see p. 135 proposed to replace the originally used description of the native proteins as "globular" by "corpuscular" since the tightly built kinetic units are often not isodiametrically dimensioned, indeed are sometimes even considerably extended in one direction, for example in some virus proteins.

[2] Only a few cases are known in which corpuscular proteins separate in the form of drops. One can in these cases merely doubt whether these are really true coacervates, that is to say, amorphous colloid-rich liquids. Compare the example discussed later on, edestin on p. 241 (this §).

[3] In some virus proteins no real crystalline phase is produced but a paracrystalline one. See on p. 243.

[4] Example: If one adds sufficient alcohol to a gelatin sol at 50°, coacervation is produced; if one then adds water at 50° the coacervate goes completely into solution again. If one allows the coacervated system to cool to room temperature and then adds cold water, the coacervate drops do indeed swell very considerably but no longer go into solution. This is attributable to the gelation of the coacervate produced by the cooling. On warming they go into solution.

[5] Crystalline arrangement of portions of the macromolecule occurs however in certain gels (see Chapter XII). For the crystallinity of fibrous protein materials of biological origin see: W. T. ASTBURY in E. BAMANN and K. MYRBÄCK, *Die Methoden der Fermentforschung*, Georg Thieme, Leipzig 1941, Bd. I.

When however by various causes, for example at higher temperatures, the previously adequate intramolecular bonds fail, *denaturation* occurs, this change consisting of a partial or total unrolling of the original tightly folded structure.

Naturally the solubility is fundamentally changed by this process, since now all the groups along the main chain of the macromolecule will determine the solubility behaviour.

When groups, originally situated in the inside of the tighly built kinetic unit, are for the greater part not favourable for solubility in the given medium [1], the protein after *denaturation* will decline greatly in solubility.

If the protein was originally in solution at its isoelectric point, then denaturation is followed by flocculation.

At other pH values it can still remain in solution but if the pH is brought to the isoelectric point or if some salt is added, flocculation again follows. Under certain circumstances "apparent single colloid systems" can be formed. Such systems of the solid type (see p. 236, § 1 d) are produced for example from the white of an egg on boiling, in which the concentrated sol is transformed into a gel.

For further questions of detail regarding denaturation and its possible reversibility reference may be made to the recent survey of the literature by ANSON [2].

Protein crystals have long fascinated many investigators by their peculiar properties. Many observations can be found in the older literature in which mention is made of volume changes in these polyhedral bodies on change of the composition of the mother liquor in which they are immersed. The observations that these volume changes may be accompanied by changes in the crystallographic angles are very remarkable. It was thus not surprising that serious doubt arose as to whether protein crystals could indeed be considered as true crystals [3]. This doubt was further increased by the fact that, after the introduction of the newly developed X-ray method, for a long time no one succeeded in obtaining X-ray diffraction patterns.

BERNAL and CROWFOOT [4] first succeeded in finding the reason for these failures. They showed with pepsin crystals that genuine X-ray diffraction patterns can be obtained when the crystals are investigated while enclosed in their own mother liquor (in the drying of the crystals customary up till then a profound change in internal structure — denaturation — can occur).

These and other results [5] show that protein crystals must really be considered as true crystals to the extent that the corpuscular protein molecules are arranged in them in a three-dimensional lattice and that the mother liquor is essential for their integrity.

---

[1] For example hydrophobic groups or groups which originally formed intramolecular salt bonds by their opposed charges and now are also available for forming intermolecular salts bonds when the pH is favourable for this.

[2] M. L. ANSON, *Denaturation and properties of protein groups*, in *Advances in Protein Chemistry*, edited by M. L. ANSON and J. T. EDSALL. Vol. II. Academic Press Inc., New York 1945.

[3] See for example a survey article by G. R. KATZ, *Die Quellung*, in *Ergebnisse der Exakten Natur- wissenschaften*, Bd. III, 1924, p. 322—325.

[4] W. T. ASTBURY, *Cold Spring Harbor Symposia on Quantitative Biology*, II (1934) 15. J. D. BERNAL and D. CROWFOOT, *Nature*, 133 (1934) 794; W. T. ASTBURY and R. LOMA, *Nature*, 133 (1934) 795.

[5] For a recent review see I. FANKUCHEN, *X-Ray Diffraction and Protein Structure*, in *Advances in Protein Chemistry*, edited by M. L. ANSON and J. T. EDSALL. Vol. II, Academic Press Inc., New York, 1945.

Now it was already known that many protein crystals contain the micro-units of the mother liquor as well as protein, for example $(NH_4)_2SO_4$ and H_2O if the protein crystallised from a sol containing $(NH_4)_2.SO_4$

Density measurements by ADAIR and ADAIR [1] lead them to the conclusion that protein crystals can be considered as phases of continuously variable composition, which contain protein and in which micro-components are also present which possess the ability to enter and to leave and thus can be substituted by others without the crystal structure being radically changed by this. Thus we arrive at the conclusion that protein crystals are true crystals as far as the lattice formed by the corpuscular macromolecules is concerned but that they differ appreciably from the usual type which one encounters normally in crystals formed exclusively of micro-units and which contain another type of matter in addition to the one considered.

In this latter case the constituents, for example the species of molecule under consideration and the extra molecules, e.g., hydrate water, present, share side by side in the construction of the crystal in well defined mutually simple molecular ratios.

Here however as a consequence of the large dimensions of the regularly arranged corpuscular protein molecules, such large interstices remain between them that a part of the mixture of micro-units filling them can be considered as falling outside the range of the real binding forces which originate in the macromolecules.

The protein crystals may thus be considered as phases, of a composition which is continuously variable within certain limits, which, when one leaves the micro-units out of consideration and considers the remaining corpuscular molecules arranged in a three-dimensional lattice, exhibit the usual pattern of a crystal.

To characterise such a phase a separate term is perhaps appropriate, for which we propose the expression "*Colloid crystal*" already used incidentally a few times.

From the standpoint of "Colloid Science" we are interested in our study objects in the first place because these contain colloids, that is to say, macromolecules. The nature of a phase containing colloid is then in the first place characterised by the relative arrangement and degree of mobility of the macromolecules present in it. On the grounds of this criterion terms such as Sols and Coacervates have indeed also become current to denote phases variable in composition within wider or narrower limits, in which, after mentally removing the micro-units still present as well in them, the macromolecules in a certain respect behave similarly to molecules in a gas or in a liquid

The colloid crystals bathed in their mother liquor are often very soft bodies and can sometimes be greatly disturbed in their external shape by shifting the cover-glass and be changed into a pool of liquid which at first sight does not differ from a coacervate. With the relatively rarified structure of the colloid crystals this is not very surprising.

In this connection the observations of HOLWERDA [2] are interesting, concerning the separation of *edestin* from its NaCl solution on dilution with water. Edestin is a plant globulin which dissolves in 10% NaCl and on dilution with water crystallises out. HOLWERDA followed this separation under the microscope, water being slowly added during the process. Initially the edestin separates in the form of drops, which

---

[1] G. S. ADAIR and M. E. ADAIR, *Proc. Roy. Soc. London B*, 120 (1936) 422.
See also M. PERUTZ, *Trans. Faraday Soc.* 42B (1946) 187.
[2] K. HOLWERDA, *Bioch. Z.*, 279 (1935), 353 see p. 355—356; 282 (1935), 317, see p. 337—338.

however on the addition of more water are gradually transformed into the typical fairly regular hexagonal shape of the edestin crystals. Triangular shapes with much rounded corners were observed in the process as transitional shapes which via hexagonal forms with three long and three short sides finally go over into the regular hexagonal form.

It is thus remarkable that this transition takes place continuously, in other words, that the hexagonal form does not occur in the drop as a small crystal and grows at the expense of the drop but the change in shape concerns one and the same phase.

The transitional shapes are again readily deformable under the cover glass by pushing it backwards and forwards under pressure and they can also unite to an apparent homogeneous coacervate layer by centrifuging (up to a certain limit).

HOLWERDA comes on the basis of this and other evidence to the view that the primary separation of the edestin is not comparable with ordinary coacervation but that forces producing order play a part right from the start. With too much NaCl in the system these are too weak to make themselves distinguishable at the periphery in a polyhedral shape.

The interfacial tension, which always attempts to produce the smallest possible surface, predominates and so spherical bodies of edestin are produced.

The ordering forces increase relatively to the interfacial tension in proportion as the NaCl concentration decreases, so that the interfacial tension is now only able to round off the sharp corners (triangular forms with rounded corners) and finally the internal order is completely successful in expressing itself in the external shape (fairly regular hexagonal shape). We can agree completely with this and are also of the opinion that the drop-like separation is not coacervation[1] but a special case of crystallisation.

As even in the world of the micro-units the production of crystals does not always provide a guarantee of the purity of the substance separated, so is this the case to a higher degree in the domain of colloid crystals. We then have as the criterion that a protein has been obtained crystalline in a pure state that the so-called phase rule test[2] should apply, in which the solubility of the preparation is investigated
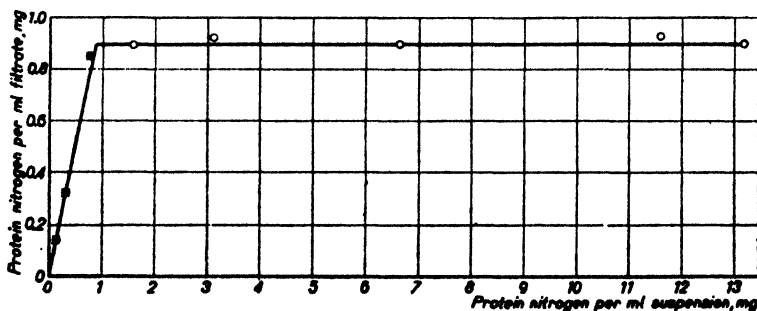
Fig. 5. Solubility of crystalline chymotrypsinogen in 0.25 saturated ammonium sulphate at 10° C in the presence of increasing quantities of solid phase (NORTHROP 1939).

---

[1] By coacervate we would understand a liquid colloid-rich phase in which there is no question of an ordering over any extensive region, that is to say, an amorphous phase. The edestin drops, according to what has been said above, do possess an ordering which is only disturbed at the periphery.

[2] J. H. NORTHROP, *Crystalline Enzymes*, Columbia Univ. Press, New York 1939.

as a function of the added amount of crystalline material. According to the phase rule the solubility must be independent of the amount of "Bodenkörper" when the protein really consists of only one component, see Fig. 5.

Just as also in the world of the micro-units phases are known which in their degree of order occupy a place in between the crystalline and the liquid, so one also finds them with some corpuscular proteins with very extended molecules. Under various conditions tobacco mosaic virus protein can be precipitated from its solution in the form of fibrous aggregates which were first looked upon as crystals but on closer investigation proved to be paracrystals.

The mutually parallel, very extended giant molecules, which themselves again probably consist of a number of subunits, together with water layers sandwiched in between form a liquid crystalline phase. In this phase the packing of the virus protein molecules is variable according to the circumstances, a fact which points to the presence of long range interparticle forces[1].

§ 3. COACERVATION AND COACERVATES

In the older literature cases of partial miscibility in particular of solutions of protamines (a class of proteins) have been described without this causing any surprise[2]. Other cases of separation into two liquid layers were gradually discovered, but they only began to attract the attention of investigators when the opinion became more and more general that the behaviour of all "Colloids" depended on boundary phenomena (p. 1 Ch. I § 1a). According to this view a sol had always to be considered as a two-phase system, in which the one phase (the disperse phase) is very finely divided in the bosom of the second (the dispersion medium, the continuous phase).

The striking fact that sols belong more or less clearly to two different types (namely those dealt with in Volume I and Volume II of this book) lead Wo. OSTWALD[3] to assume that the striking differences between these two types of sols were based on the solid and the liquid nature respectively of the dispersed phase. OSTWALD based on this his classification of the sols into "Suspensoids" and "Emulsoids" and he saw in the partial miscibility phenomena which occur in the latter a direct proof of his conception.

Indeed other examples, such as gelatin or casein sols + sulphates[4], gelatin + sulphosalicylic acid[5] also belong to the type of sol which he had called the emulsoid type.

[1] J. D. BERNAL and I. FANKUCHEN, Nature, 139 (1937) 923. See further recent reviews: I. FANKUCHEN, X-Ray Diffraction and Protein Structure, in Advances in Protein Chemistry, edited by M. L. ANSON and J. T. EDSALL, Vol. II, 1945, Academic Press Inc. New York and N. W. PIRIE, Physical and Chemical Properties of Tomato Bushy Stunt Virus and the Strains of Tobacco Mosaic Virus in Advances in Enzymology, edited by F. F. NORD and C. H. WERKMAN, Vol. V, 1945, Interscience Publishers New York. See also Volume I of "Colloid Science" for long range forces.
[2] A. KOSSEL, The Protamines and Histones, Monographs on Biochemistry, Longmans Green and Co. London, New York, Toronto 1928, compare p. 26 salmin sulphate: p. 30 clupein sulphate; p. 33 scombrin sulphate; p. 38 sturin sulphate.
[3] Wo. OSTWALD, Kolloid-Z., 11 (1912) 230.
[4] K. SPIRO, Beitr. Chem. Physiol. u. Path., 4 (1903) 300; W. PAULI, Kolloidchem. Beihefte, 3 (1912) 382—383.
[5] Wo. OSTWALD, Kolloid-Z., 43 (1927) 131.

The idea was therefore that in the original sol (for example the gelatin sol) the sol particles, united with a certain quantum of water, are present as ultramicroscopic liquid drops and that by the action of, for example, sulphates, these drops coalesce to microscopically visible drops or even to a macro liquid layer. In this line of thought the visible separation into two liquid layers of an emulsoid sol is only a change in the degree of dispersion of a second liquid phase already present in the original sol.

It did in fact cause OSTWALD no surprise, but rather strengthened him in his ideas when it was found that the phase rule[1] does not hold for the partial miscibility in the system gelatin + H_2O + sulphosalicylic acid.

a   b
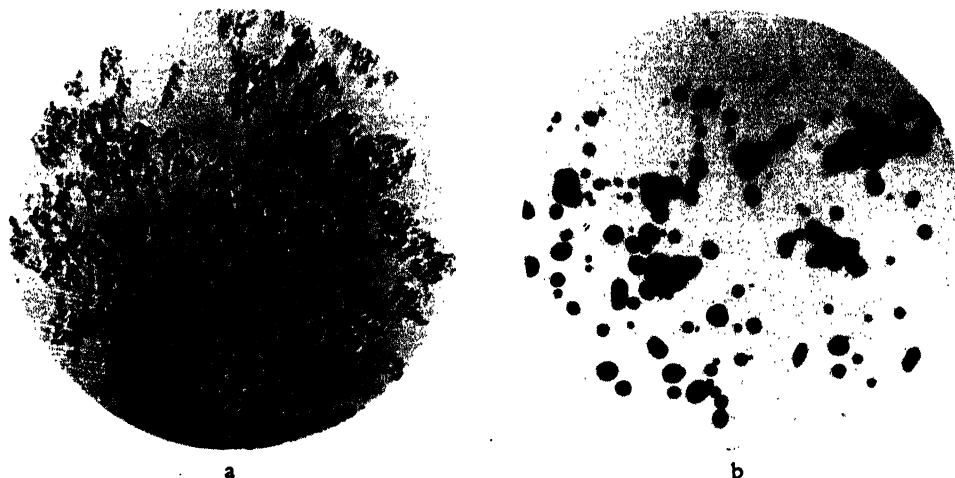
Fig. 6.   Flocculation or coacervation with Na nucleate + hexol nitrate.
a: At room temperature flocculation is produced. 140 × lin.
b: After heating of the system reproduced in a) the floccules have coalesced into coacervate drops and on cooling the latter have weakly vacuolised. 115 × lin.

BUNGENBERG DE JONG and KRUYT[2] in a preliminary communication on partial miscibility phenomena in "hydrophilic" sols also proceeded from the conception that these sols are two-phase systems. They were also of the opinion that the visible separation into two liquid layers is not accompanied by a change in the number of phases. Since the term "partial miscibility" already had a well defined meaning (separation of one phase into two coexisting phases) it appeared desirable not to use the same term for the formation of two liquid layers in "hydrophilic" sols. Consequently they introduced the term *coacervation* (from the latin *acervus* = aggregation, heap and the prefix *co* to signify the preceding union of the colloidal particles). Coacervates are colloid-rich liquids which are not spontaneously birefringent.

By means of a great number of fresh cases, they were able to show that the causes which call forth coacervation sometimes bring about flocculation if the con-

---

[1] See however for the non-validity of this argument small print on p. 253.
[2] H. G. BUNGENBERG DE JONG and H. R. KRUYT, *Proc. Koninkl. Nederland. Akad. Wetenschap.*, Amsterdam, 32 (1929) 849; *Kolloid-Z.*, 50 (1930) 39.

ditions are changed but relatively little. Thus Na nucleate gives flocculation with hexol nitrate at room temperature and coacervation at $40°$ [1] (see Fig. 6). Again on comparison of very similar systems, in the one case coacervation, in the other flocculation takes place under the same conditions. Compare the flocculation of nucleate with hexol nitrate at room temperature and the definitely liquid coacervation of gum arabic with hexol nitrate also at room temperature. See Fig. 35 in Chapter X on p. 385.

Coacervation and flocculation are thus very closely related phenomena and therefore it seemed natural that coacervation could in principle also be explained by an application of the "Stability Theory" previously developed by them [2]. According to this theory hydrophilic sols are characterised by two stability factors: capillary electric charge and hydration.

Coacervation would thus demand a removal of the two stability factors. In fact on studying the changes which $\eta_{sp} = \dfrac{\eta_s - \eta_o}{\eta_o}$ of the sol undergoes when the conditions are gradually realised which lead to coacervation, results were obtained (considerable decrease of n_{sp}) which could be interpreted in that sense. The essential change of the sol particles in these preliminary processes, without which no coacervation sets in, were interpreted in that manner as a partial dehydration.

The original sol particle — idealised for the sake of simplicity to a massive sphere — was considered to be surrounded by a hydration coating of considerable size, in which the water was bound less and less tightly towards the periphery. Such a sol particle would owe its stability just to this diffuse character of the solvate coating, since the latter is not sharply defined at its periphery and consequently also possesses no free surface energy. In fact if it possessed a free surface energy then the sol particles would unite with one another through their solvate coatings on collision.

The preparative process, which enables the union of particles into a coacervate, is based then according to this view on the transformation of the diffuse solvate coating into one which really has a sufficiently concrete outer boundary.

Now according as the solvation (and possibly discharge) sets in and as a result of it the thickness of the solvate coating decreases, its delimitation would become more and more concrete at its periphery.

It was thus supposed that a coacervate consists of colloidal particles which have merely united through their concrete solvate coatings. See Fig. 7.

Since the actual particle nuclei are still always displaceable with respect to each other the coacervate has the nature of a liquid and in the cases investigated POISEUILLE's law was found to hold very accurately (Newtonian liquid [3]).

On further change of the variables which lead to coacervation, the desolvation process will at first continue in the same direction as a result of which the thickness of the solvate coating separating the particles decreases, that is to say the coacervate becomes poorer in solvate [4], i.e., richer in colloid.

---

[1] This example is not taken from the above mentioned publication but from a later one (H. G. BUNGENBERG DE JONG and F. A. MENALDA, *Biochem. Z.*, 257 (1933) 62), where it was also found that the transition in question from floccules into coacervate drops can also occur at room temperature on the addition of 15% resorcinol.

[2] See Chapter VII § 4, p. 197.

[3] H. G. BUNGENBERG DE JONG, H. R. KRUYT and J. LENS, *Kolloidchem. Beihefte*, 36 (1932) 429, See p. 439.

[4] Under these circumstances coacervate drops exhibit vacuolisation. Compare p. 443, Ch. XI § 2.

If one brings about a very considerable desolvation of the particles all at once, then the coacervate will become so extremely viscous on account of the appreciably smaller thickness of the solvate coatings separating the particles that the fusion of small drops into large ones can occur only extremely slowly. Morphologically the coacervation then assumes the character of a flocculation.
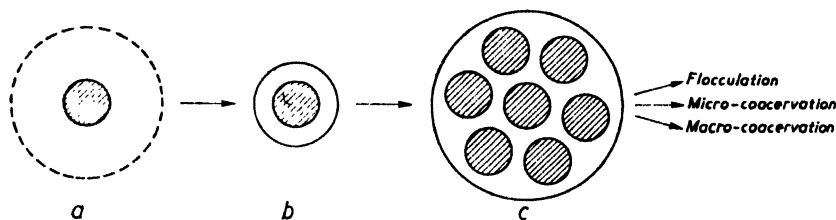
Fig. 7. Original scheme for the mechanism of coacervation.
a: Sol particles with diffuse solvate coating.
b: Particles with concrete solvate coating (after desolvation).
c: The particles fuse into a coacervate with their concrete solvate coating. The concrete boundary is afterwards only present at the boundary coacervate-equilibrium liquid.

The above sketched theory of coacervation served as a useful guide in the further experimental work on the partial miscibility phenomena in "hydrophilic" sols[1] and after its fundamental assumptions had later become more and more doubtful and another interpretation had already been put forward it still served at least as a service-able scheme for the preliminary recording of the experimental facts[2].

Now however it has only historical value, since we are now convinced that the starting point — the two-phase nature of the original sol and the stability theory based on it — can no longer be maintained.

From the macromolecular standpoint the original sol is merely the true solution of the colloid and thermodynamically the coacervation is just as much a true partial miscibility in the sense of the phase theory as the partial miscibility which occurs in systems which consist exclusively of micro-units (for example phenol-water; alcohol-water-$(NH_4)_2SO_4$).

Nevertheless it is a very extreme case of partial miscibility in so far as the two phases differ practically only in composition through the colloid component being present in the one phase (the *coacervate*) in high concentration and in contrast in very low concentration in the other phase (the *equilibrium liquid*). If besides the colloid there is, for example, also a mixture of two kinds of micro-units present, then their relative poportion in coacervate and equilibrium liquid is more or less the same (for more detailed discussion and restriction of this statement see below p. 252 and p. 254).

As this partial miscibility is in the first place characterised by the distribution of one of the components throughout two liquid phases and this one is just the colloid

[1] See reviews of Coacervation:
H. G. BUNGENBERG DE JONG, *Protoplasma*, 15 (1932) 110. H. G. BUNGENBERG DE JONG, *La coacervation, les coacervats et leur importance en biologie*, Tome I en II, Hermann et Cie, Paris 1936.
[2] See review of Coacervation: H. G. BUNGENBERG DE JONG, *Kolloid-Z.*, 79 (1937) 233, 334; 80 (1937) 221, 350.

component, from the standpoint of Colloid Science it serves a useful purpose to denote this special case of true partial miscibility by a specific technical term. It is for this reason that we retain the terms *coacervation, coacervate* and *equilibrium liquid* in what follows. We desire further expressly to restrict the coacervation concept to the formation of an *amorphous* (not spontaneously birefringent) colloid-rich liquid.

When using these terms we must however completely sever ourselves from the original coacervation theory as this is expressed in the symbols of Fig. 7.

First of all "sol stability", that is to say, solubility of the colloid, is not based on the acquirement of a fairly thick solvate coating. The very high values of $[\eta] = \dfrac{\eta_s - \eta_0}{c \cdot \eta_0}$, with gelatin, agar and other biocolloids, from which it was concluded that thick solvate coatings exist around the particles, are based according to our modern conception on the loose skein shape of the "linear" macromolecule (p. 210 Ch. VII § 6).

With the corpuscular protein, egg albumin, $[\eta]$ is low and nevertheless it forms completely stable solutions. The solubility of macromolecular colloids thus cannot depend on the acquirement of thick solvate coatings and with this the starting point of the original coacervation theory is nul and void.

Before attempting to give a new interpretation of the typical coacervation of macromolecular colloids in the following pages, we wish also to point out that no good examples are yet known to us in which a genuinely liquid coacervation of corpuscular proteins occurs [1,2]. Under circumstances in which this could be expected, for example with $(NH_4)_2 SO_4$, either crystallisation or a so-called amorphous flocculation occurs.

Microscopic examination of this amorphous flocculation provides no decision on the nature of the separated protein-rich phase. Naturally it is possible that this consists of a coacervate but in that case it is certainly very poor in water and has much rather the character of an amorphous solid substance.

Examples of typical coacervation are to be found among the "linear" macromolecular colloids especially among those which belong to the highly viscous type [3]. This is an indication that it will be important for the theory of a typical coacervation to associate it with the "skein" shape of the macromolecule. With this disappears all serviceability of the scheme of Fig. 7 for an explanation of a typical coacervation in view of the fact that it is not only incorrect as regards the symbolisation of the hydration but also as regards the conception of the colloidal particles as solid spheres.

Only the "amorphous" flocculation of corpuscular proteins could still be represented by Fig. 7 in so far as the separated highly dispersed colloid-rich phase is not microcrystalline or paracrystalline, if at any rate one replaces the very great hydration by a suitably smaller one, similarly the spheres by appropriately shaped corpuscles and thinks of these as placed in the "coacervate" at short distances from each other but in such a way that there is still no question of a rigorous order

---

[1] Genuinely liquid coacervates can indeed be produced in which a macromolecular colloid of the linear type takes part besides a corpuscular protein, for example, the complex-coacervation of serum albumin (positive) — gum arabic (negative). See p. 233, Fig. 2.

[2] As regards the so-called coacervation of edestin, see p. 241 § 2.

[3] Coacervation can also occur with association colloids, for example, oleate+KCl. As a preliminary to the coacervation high viscous systems occur in a certain range of KCl concentrations, at higher KCl concentrations the viscosity falls and at still higher coacervation sets in. See p. 701 and 708 Ch. XIV § 4 and 5.

over larger distances. The chance of the production of an order, especially when the corpuscular macromolecules are very oblong, will however be great then and if it does occur over larger distances, we would prefer not to speak of a coacervate but of a paracrystalline or crystalline phase.

§ 4.   COACERVATION OF MACROMOLECULAR COLLOIDS OF THE HIGH VISCOUS TYPE FROM MODERN POINTS OF VIEW

The concept of the statistically kinked macromolecule can be of service in interpreting in another, more satisfactory way the facts on which the old coacervation theory was based.

These facts are:

1.   If one gradually changes the composition, for example by adding a micromolecular substance, say Na_2SO_4, of a dilute sol of the high viscous type, for example isoelectric gelatin, then the expression $\frac{\eta_s - \eta_o}{\eta_o}$ decreases sharply in a certain concentration range of the added substance, previous to the coacervation (p. 218 Fig. 28).

2.   On just exceeding the coacervation limit the coacervate frequently still contains a relatively large amount of water + micro-units. This content at first decreases further on further addition of the micromolecular substance.

If now one bears in mind that the colloid in question (for example gelatin) belongs to the high viscous type, then if sufficient soluble groups are present along the linear macromolecule (for example oxyproline built into the gelatin molecule) these macromolecules will occur in dilute sols in the state of the statistically kinked macromolecule [1].

If now by the addition of the micromolecular substance one alters the solvent in that direction in which the macromolecule no longer remains soluble (gelatin is insoluble in conc. Na_2SO_4 or alcohol), it will gradually be situated in more and more unfavourable conditions for solution. The affinity of the various groups along the macromolecule for the solvent decreases, arrives finally at the same order of magnitude as their mutual affinity and also decreases still further below this.

The result will be, that the macromolecule, which, in water and in dilute solution of the added micromolecular substance, is present to begin with more or less as a statistically kinked macromolecule, must begin to lose this latter character on increasing the concentration of the added micromolecular substance. Contacts of shorter or longer duration will have to begin to occur between loops of the macromolecule at the positions of the groups which first become no longer completely "soluble". As a consequence of these intramolecular contacts the circumscribed volume over which the macromolecule extends must decrease considerably. This gives an explanation of the large decrease of n_{sp} before the actual coacervation begins.

On further increase in concentration of the added micromolecular substance the changes in the state of the macromolecule will proceed still further in the direction mentioned as a result of which sufficient points of contact of more or less long duration can also be formed between loops of different macromolecules. If this intermolecular

---

[1] In the following paragraphs we set aside possible mutual interactions already present between the peptide groups or between groups in the side chains of the macromolecule and are thus knowingly dealing with a simplified version of the matter.

association is sufficiently great, coacervation takes place. Now also on further increase of the concentration of the added substance the changes will continue in the same direction: more intramolecular points of contact and more intermolecular ones are produced; as a result of which point (2) becomes understandable.

When we summarise the above statements, the sharp fall in n_{sp} preceding the coacervation is not attributable to dehydration but to a great reduction of the amount of occlusion-liquid inside the macromolecule (see already p. 209, Ch. 7 § 6).

Further the coacervate is to be regarded as an association of macromolecules, in which the points of contact must not be thought of as static but dynamic, since it is still a typical[1] liquid, though a viscous one. These conceptions, at which we have thus now arrived, stand diametrically opposed to the original ideas, in which it was still believed that the "particles" were permanently separated by a very considerable layer of hydration water. In the new ideas there is no longer any place for the view that the water present in the coacervate is bound as a whole by hydration forces.

By far the larger part (at least if the water content of the coacervate is not very small) is to be regarded as occlusion-water, included between the loops of macromolecular skeins associated together.

If we want to symbolize the points discussed above by simple schematic diagrams we obtain Fig. 8, which then takes the place of Fig. 7.

Fig. 8. Modified scheme of coacervation.
a: Circumscribed volume (cross section simplified to circular) of a loosely built macromolecular skein in the original sol.
b: The macromolecular skeins have become more dense and tend to associate together (here also only the circumscribed volume is drawn).
c: In the coacervate there are mutually associated macromolecules which penetrate each other at any rate with their peripheral loops. The polygons drawn do not here represent the separate circumscribed volumes. These latter penetrate each other more or less and this penetration is just the strongest at the positions of the sides of the polygon.

There is no longer any symbolisation of the hydration in these schematic pictures. Fig. 8 a then represents the circumscribed volume of the skein shaped macromolecule in the original extremely dilute sol. This circumscribed volume is drawn as spherical, although the statistically most probable shape of the macromolecular skein is actually a general ellipsoid.

The transition a → b represents the preliminary stage to the coacervation. Intramolecular associations occur as a result of which the area of the macromolecule decreases. It is probable that intermolecular associations already take place here, but these aggregates still remain in their entirety in solution.

On further change of the variable in the direction of decreasing solubility the intermolecular associations become so extensive that they are no longer soluble. So

---

[1] The coacervates, which have been investigated as regards their viscosity behaviour, behave as Newtonian liquids. See note 3 on page 245.

we arrive at the formation of a separate colloid-rich phase (the coacervate) in equilibrium with a colloid-poor phase (the equilibrium liquid).

In schematic drawing c the internal state of the coacervate is symbolized. The macromolecular skeins are here drawn as hexagonal fields fitting together, although more or less appreciable interpenetration of the contiguous skeins will take place.

If in imagination we remove all the micro-units, then the coacervated system consisting of two coexisting liquids assumes the character of a condensation equilibrium: consisting of a colloid liquid (the coacervate) and a colloid vapour (the equilibrium liquid). See p. 12 Ch. I § 4 c 2.

§ 5. DIVISION OF COACERVATION INTO TWO MAIN TYPES

Coacervation can be brought about in very different ways (for example, a change in temperature, a change of pH, addition of a micromolecular substance, or of a second macromolecular substance) but we can definitely subdivide them into two large groups, according to whether the ionised groups of the macromolecule play an active part in it or not. In terms of the considerations given in § 4) this means that the reduction in solubility and the associations deriving from it are in the first case — *Simple coacervation* — concerned with non-ionised groups in the macromolecule, in the latter case — *Complex coacervation* (in the extended sense) — it are just the charges on the macromolecule which are concerned in the matter, with the formation of salt bonds.

We must still point out that the given division is not an ideal one in so far as only one mechanism is placed in the foreground each time. Cases are known in which both are in action at the same time and we could in those cases speak of mixed types. In some cases one of the mechanisms is well to the fore, for example in the coacervation of isoelectric gelatin with alcohol, we are mainly concerned with simple coacervation, although there are indications that interaction between charges of opposite sign also plays a part to a slight extent. (See p. 217 Ch. VII § 9).

In other cases both mechanisms clearly act simultaneously, for example in the coacervation of gum arabic + electrolyte + alcohol (see p. 396, 400 Ch. X § 3f, 3g).

In this case the alcohol has a two-fold action, on the one hand it reduces the solubility of the non-ionised groups, on the other hand its strengthens the interaction between the cations of the added salt and the negative ionised groups (carboxyl groups) of the gum arabic.

Furthermore we would point out that the main division has nothing to do with the number of colloid components present.

Cases are known in both main groups in which coacervation occurs when two colloids are present as well as when only one colloid is present.

Again in the mixture of the same two colloids simple or complex coacervation can appear according to the circumstances, a striking example of which we have still to discuss. See p. 255 § 8.

For the further discussion of coacervation in this Chapter we shall restrict ourselves to some examples of simple coacervation.

Complex coacervation is dealt with more extensively in the Chapter, "Complex Colloid Systems", Ch. X p. 335. For a brief characterisation see however the right-hand column of the summary on p. 256.

§ 6. SIMPLE COACERVATION OF ISOELECTRIC GELATIN WITH ALCOHOL OR NA₂SO₄ [1]

If we add alcohol gradually to an isoelectric gelatin sol (to avoid complications by gelation the temperature should be sufficiently high, say 50° C) then the mixture remains clear up to a certain alcohol concentration. Turbidity is produced on further addition and on microscopic examination the presence of a large number of coacervate drops is established which (at 50°) readily coalesce with each other. If the alcohol concentration is not too high, a coherent layer of coacervate is produced after some time. If however one has added more alcohol, then coalescence and formation of a coherent coacervate layer takes place with greater difficulty on account of the higher viscosity of the coacervate. At a sufficiently high alcohol concentration, lumpy masses or cohering masses of floccules are finally obtained.

If in the last two cases one adds water carefully (lowering of the alcohol concentration) the coacervate again becomes less viscous and the coalescence to a coacervate layer again becomes easier.

From what is stated above it follows that coacervation does not represent a temporary intermediate state between sol and flocculation. For the case in which a flocculation has the nature of a coacervate, on the contrary the opposite holds: The flocculate is transformed into a coherent coacervate layer after a sufficiently long time (See p. 234 § 1 c).

The coacervation is completely reversible: the coacervated system passes again into the sol state when one lowers the alcohol concentration sufficiently by adding water. Furthermore the composition of coacervate and equilibrium liquid is completely determined at a given total composition of the system. Coacervate layer and equilibrium liquid are in thermodynamic equilibrium; the way by which one arrives at any *particular final composition* of the total system is of no influence on the composition of coacervate and equilibrium liquid.

Fig. 9 reproduces schematically in a so-called ternary diagram the analytical results [2] for the system
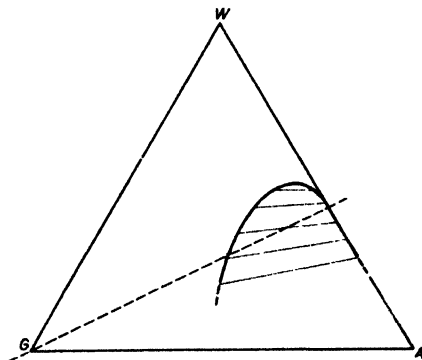
Fig. 9. Coacervation in the system isoelectric gelatin (G) + water (W) + alcohol (A) (schematic). The coacervates lie on the branch of the curve which proceeds into the plane of the triangle, the equilibrium liquids on the branch of the curve which is adjacent to the side WA. The tie lines are intersected by any straight line through the corner G (e.g. the dotted one) in such a way that the ratio alcohol: water in the coacervate is smaller than that in the equilibrium liquid.

isoelectric gelatin (G) — water (W) — alcohol (A).

---

[1] L. W. J. HOLLEMAN, H. G. BUNGENBERG DE JONG and R. S. TJADEN MODDERMAN, *Kolloid-chem. Beihefte*, 39 (1934) 334.

[2] The analytical results shown schematically in Fig. 9 are valid for the coacervates and equilibrium liquids produced from systems which contain almost the same percentage of G in the total system. (30 g gelatin dissolved in 100 g water and diluted with 470 g of an alcohol-water mixture of varying composition). The total mixtures then contain 5% gelatin. For the significance of these see the small type on p. 253.

The region in which partial miscibility occurs is bounded by a curve, one branch of which lies close alongside the W.A. side. This branch represents the compositions of the equilibrium liquids (p. 246). The coacervates are to be found on the other branch which proceeds towards the middle of the plane of the triangle. The region of partial miscibility is only partially represented since at higher alcohol concentrations the nature of the coacervate (lumpy or flaky masses on account of high viscosity) no longer makes it possible to realise the separation into homogeneous liquid layers required for the analysis.

It is seen from the figure that the coacervate contains alcohol as well as gelatin and water.

It is now of importance to us to consider the direction of the tie lines which join coexisting phases, in this case coacervate and the corresponding equilibrium liquid.

In the ideas concerning the internal state of the coacervate (set out in § 4 p. 248) we have shown that the water (and possibly other micro-units present) must in no case be regarded in its whole as hydration (solvation) but for the greater part as occlusion liquid.

If all the water and alcohol present in the coacervate were occlusion liquid, then one would have to expect that the tie lines would run in the direction of the corner G.

In fact a line drawn through the corner G (for example the dotted one) indicates the geometrical position of all compositions of the three components for which the ratio of A and W is constant. It is seen from the figure that the tie lines indeed run roughly but not precisely in the direction of the corner G.

A line through G (for example the dotted one) intersects the tie lines in such a way throughout that the A + W mixture present in the coacervate (which of course also contains G) contains relatively more water and less alcohol than the A + W mixture in the coexisting equilibrium liquid.

This needs cause no surprise since the prerequisite for a more or less extensive unfolding of the macromolecule into skeins, in which space is available for occlusion liquid, is that certain groups along the macromolecule are in the dissolved state, that is to say are solvated. Now gelatin is soluble in water, insoluble in alcohol, so that water is what is necessary for this required solvation.

One can thus consider the water + alcohol present in the coacervate as consisting of two parts namely of an alcohol + water mixture of the same composition as that present in the equilibrium liquid and to be regarded exclusively as occlusion liquid, and a certain surplus of water which is be regarded as true hydration of the gelatin. From the analytical data one calculates in this way a surplus of water of about 0.3 g per gram of gelatin, a figure which is not improbable as far as order of magnitude is concerned.

The coacervation of isoelectric gelatin with Na_2SO_4 has also been investigated in detail and the results obtained fit very well with those obtained in the coacervation with alcohol. See Fig. 10. In this case also the coacervate contains Na_2SO_4 as well as water and the concentration of the Na_2SO_4 in this mixture is smaller than in the equilibrium liquid (the tie lines are intersected by a line through the corner G in the ternary diagram in the same way as in Fig. 9).

If one again regards this $Na_2SO_4 + H_2O$ present in the coacervate as occlu-

sion liquid + surplus water attached to gel-
atin, somewhat higher figures are obtained
in this case for the hydration than in the
case of the coacervation with alcohol and
furthermore these figures fall with increas-
ing Na_2SO_4 concentration (1.1 → 0.7 g
per gram of gelatin).

Fig. 10. Coacervation in the system isoelectric
gelatin (G) + water (W) + Na_2SO_4 (S) (schematic).
The coacervates lie on the branch of the curve
which proceeds into the plane of the triangle, the
equilibrium liquids on the branch of the curve
which is adjacent to the side W S. The tie lines are
intersected by any straight line (e.g. the dotted one) through the corner G in such a way that the
ratio sulphate: water in the coacervate is smaller than that in the equilibrium liquid.
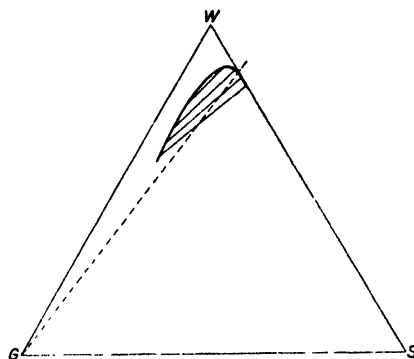
        The use of the ternary diagrams is in reality only permissible if 1) one assumes that the phase
rule holds also for phase equilibria in which macromolecular colloids take part and 2) when the
number of components is in fact three.
        As far as 1) is concerned we are of the opinion that there is no need to doubt it in view of the
modern ideas of sols of macromolecular colloids as true solutions and further in view of the character
of the thermodynamic equilibrium between coacervate and equilibrium liquid.
        As regards 2) with gelatin we immediately are faced with the fact that it is really not one component
in the sense of the phase theory but a mixture of many, for the rest closely related components of,
for example, different chain lengths. When thus in the above paragraphs the mixtures of gelatin +
water + alcohol or gelatin + water + Na_2SO_4 were considered as ternary systems, this is a conscious
simplification. It is then not surprising that a complete description is not possible in such a diagram.
If the system G + W + Na_2SO_4 were really a ternary system, then the position of the partial
miscibility curve in the triangular diagram would have to be independent of the total amount of
gelatin in the total system. That is not the case according to an investigation specially directed to this
purpose, for the details of which reference may be made to the original publication [1].
Since it follows from this that the position of the partial miscibility curve in Fig. 9 and 10 still
depends on the total amount of gelatin in the total system (Fig. 9 holds for 5% G in the total system;
(see note 2 on p. 251), it gives the impression that the phase rule does not hold in principle.
This conclusion, however, looses all foundation when one remembers that the gelatin is not
one component but a mixture of many, for the rest related components.
        The non-validity of the phase rule is thus only apparent and rests on an underestimation of
the real number of components.
        We shall meet another example of underestimation of the number of components in complex
coacervation (see p. 366 Ch. X § 2 m).

§ 7. SIMPLE COACERVATION OF ISOELECTRIC GELATIN WITH RESORCINOL

        In the immediate neighbourhood of the isoelectric point coacervation is brought
about by phenols (phenol, the three diphenols as well as the three triphenols, d-
catechol, chebulic acid, digalloyl glucose and tannin) [2].

    [1] L. J. W. HOLLEMAN, H. G. BUNGENBERG DE JONG and R. S. TJADEN MODDERMAN, Kolloid-
chem. Beihefte, 39 (1934) 334; see p. 399—406.
    [2] H. G. BUNGENBERG DE JONG, Rec. trav. chim., 48 (1929) 494.

The combination isoelectric gelatin (G) + water (W) + resorcinol (R), the analytical results of which are depicted schematically in Fig. 11, has been investigated in more detail [1].

The coacervates are to be found on the loop lying in the plane of the triangle, while the equilibrium liquids on account of their very slight G content are to be found practically on the side W-R.

Compared with coacervation with alcohol (or with Na_2SO_4) the following points of difference appear:

1. small resorcinol concentrations are already sufficient for coacervation.
2. coacervation ceases completely at sufficiently high resorcinol concentrations.
3. the slope of the tie lines is just the opposite with respect to a line through G (dotted line in Fig. 11).
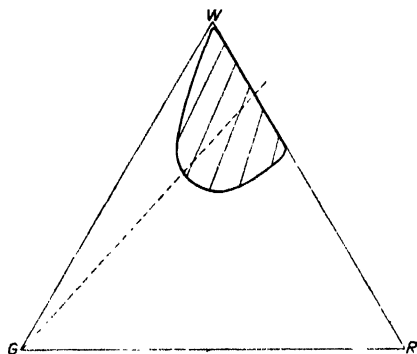
Fig. 11. Coacervation in the system isoelectric gelatin (G) + water (W) + resorcinol (R).
The coacervation region is here represented by a closed curve.
The coacervates lie on the arched branch of the curve which runs inside the plane of the triangle, the equilibrium liquids on the branch of the curve which runs close to the side W R.
The tie lines are intersected by any straight line (e.g. the dotted one) through the corner G in such a way that the ratio resorcinol: water in the coacervate is greater than that in the equilibrium liquid.

From measurements of viscosity it was found, that η_{sp} decreases considerably in this case also previous to or simultaneously with the production of coacervation, but increases again considerably at high resorcinol concentrations at which coacervation ceases again (see p. 201, Fig. 13 b).

The picture which we have made in the previous pages of the establishment of coacervation is thus also applicable again in this case. Resorcinol in small concentrations effects the production of intra and intermolecular contacts by which the macromolecular skeins begin to assume a smaller circumscribed volume and begin to associate to an increasing degree, as a result of which phase separation into a coacervate layer and an equilibrium layer containing practically no gelatin is established.

On further increase of concentration the contraction of the associated macromolecular skein at first still increases but now on still further increase of the resorcinol concentration changes occur in the opposite sense, a feature not present in the coacervation with alcohol or Na_2SO_4.

The finer mechanism by which the intra and intermolecular contacts are established must indeed be quite different and the above mentioned points 2. and 3. give an indication in that direction.

According to 3. the water-resorcinol mixture present in the coacervate is *richer* in resorcinol than the coexisting equilibrium liquid. One can therefore again consider this resorcinol + water mixture present in the coacervate as consisting of occlusion liquid of the same composition as the equilibrium liquid + a surplus. This surplus

---

[1] See note 1, p. 253.

is now not water (as in the coacervation with alcohol or Na_2SO_4) but just the substance which was added, i.e., the resorcinol. This resorcinol surplus calculated per gram of gelatin in the coacervate also increases continuously in magnitude (at 4.2% resorcinol in the equilibrium liquid this amounts to 0.2 g, at 17.5% it is 0.5 g and at 30.4% it amounts to 0.9 g per gram of gelatin).

There is every reason here for speaking of a solvation of the gelatin by resorcinol which increases with increasing resorcinol concentration in the equilibrium liquid. This solvation by resorcinol m··st be held responsible for the fact that gelatin dissolves in the most concentrated resorcinol solutions. A similar dissolving power for gelatin and other proteins is also to be found in phenol made liquid by a small percentage of water.

We still have to form a picture of how resorcinol can bring about coacervation at low concentration but loses this power at high concentration.

In this one might start from the hypothesis that the resorcinol can attach itself by its phenol groups to certain polar groups of the gelatin molecule. As a polyphenol then, provided that all of the relevant groups of the gelatin are not yet occupied, it can temporarily cement together with its two phenol groups loops of one and the same macromolecule or of two neighbouring macromolecules. But then the conditions for coacervation are fulfilled.

With further attachment, if finally practically all the groups of the gelatin are occupied by resorcinol molecules, there will be almost no opportunity for this any more. The result is then that coacervation no longer takes place at very high resorcinol concentrations.

The very strong effect of tannin, a substance with a very large number of phenol groups, is understandable from the above picture.

We encounter difficulties however when we turn to phenol itself. Here only one phenol group is present and one attached phenol molecule therefore cannot bring about intra or intermolecular cementing. Here we must assume that cementing can nevertheless occur by association of the hydrophobic benzene nuclei of two attached phenol molecules. Unpublished investigations on the effectiveness of phenol, di and triphenols have in fact shown that phenol occupies a position apart. Its effect at smaller concentrations is much weaker than that of the other phenols but at higher concentrations it is indeed much more powerful. As regards the question of on what positions in the macromolecule the phenol group is attached, some indication is given by the fact that at sufficiently high resorcinol concentrations the power of gelation of the coacervates is lost.

Since we now consider it probable that the peptide groups are directly concerned in the gelation, we think we may conclude from this that the peptide bond is the position sought.

§ 8. SIMPLE COACERVATION IN MIXTURES OF CONCENTRATED GELATIN AND GUM ARABIC SOLS

In all the examples of simple coacervation so far considered, this is brought about by an added micromolecular substance (alcohol, Na_2SO_4, resorcinol). Cases also exist in which a second macromolecular substance brings about the separation into two liquid layers. OSTWALD and R. H. HERTEL [1] described this phenomenon in gelatin + H_2O + starch. If two concentrated sols are mixed, then two liquid layers form; of which each in the main contains only one of the two colloids. Charge effects play no noticeable role in this partial miscibility and the effect on it of neutral salts merely brings to light the well-known lyotropic sequences of ions. The principal requisite for the production of the partial miscibility is in this case according to OSTWALD that the two colloids should show a sufficiently great difference as regards

[1] Wo. OSTWALD and R. H. HERTEL, Kolloid-Z., 47 (1929) 158, 357.

their power of water attraction and that the colloid which has the greatest affinity for water (in this case gelatin) should be present in high concentration.

Our own experience relates to mixtures of concentrated gelatin and gum arabic sols, in which a similar partial miscibility phenomenon occurs [1].

All the properties of this partial miscibility indicate that charge effects play no part in the process and that the main requisite for the occurrence of coacervation is a water deficit in the total system (see summary, left hand column). This example is furthermore especially interesting because another type of coacervation is also possible in mixtures of gelatin and gum arabic, the so-called complex coacervation (in the narrower sense).

To what extent the conditions for the appearance of both kinds of coacervation are different and the properties of the coacervates likewise differ, follows from a comparison of the two columns of the following summary: See also the schemes as represented in Figs. 12 a and b. Complex coacervation is discussed in greater detail in Chapter X § 2, p. 338.

Simple coacervation on mixing gelatin and gum arabic sols.	*Complex coacervation* on mixing gelatin and gum arabic sols.
Also occurs at pH $>$ I.E.P. of the gelatin (at which both colloids are negatively charged).	Only occurs at pH $<$ I.E.P. of the gelatin (at which the two colloids have opposite charges.)
Is only possible on mixing concentrated sols. The coacervation disappears on adding water.	Does not occur in concentrated sols, occurs however after dilution with water. Still occurs even on mixing 0.001% sols.
Added indifferent salts do not suppress coacervation, appear rather to promote coacervation in certain cases. The ions arrange themselves in their effectiveness in lyotropic series.	Indifferent salts suppress coacervation. The valency of both ions is of primary importance in the matter. The position of the ions in the lyotropic series is of very minor significance.
The drops exhibit no disintegration phenomena in a D.C. electric field.	The coacervate drops exhibit disintegration phenomena in a D.C. electric field.
Both liquid layers are rich in colloid, each layer containing in the main one of the colloids.	One layer, the coacervate, is rich in colloids, the second layer however poor in colloid. Both layers contain gelatin and gum arabic in not far different ratios.
Principal Condition Water deficit in the total system.	*Principal Condition* Adequate charge opposition between the two colloids.

[1] H. G. BUNGENBERG DE JONG and J. LENS, *Kolloid-Z.*, 58 (1932) 209. For rectification of some examples mentioned in this, which later were found to belong to a different type of coacervation, see H. G. BUNGENBERG DE JONG, *Kolloid-Z.*, 79 (1937) 223, see p. 227.

The partial miscibility phenomena in mixtures of concentrated gelatin and. gum arabic sols raise a number of problems for which we cannot yet give any satisfactory explanation.

First of all the question presents itself, are both colloid-rich liquid layers to be considered as coacervates or is only one of them a coacervate and the other a sol? This question is very difficult to answer since with these concentrated systems almost no criteria are present for distinguishing coacervate from sol. With regard to certain wetting phenomena the gelatin-rich layer behaves as a coacervate, the gum arabic-rich layer as the corresponding gelatin-poor equilibrium liquid[1].

Fig. 12. Schemes showing the principles of two types of coacervation in mixtures of gelatin and gum arabic sols

a. *Simple coacervation* in concentrated sol mixtures. Tie lines connect liquids which are rich in G and relatively poor in A (on the left hand branch) with liquids which are rich in A and relatively poor in G (on the right-hand branch).

b. *Complex coacervation* in dilute sol mixtures at such a pH that the gelatin is positively charged, and the gum arabic has not yet lost its negative charge.

The coacervates which are rich both in G and in A are to be found on the arched branch of the curve in the plane of the triangle. The equilibrium liquids which are poor both in G and in A lie on a branch of the curve close to the water-corner of the triangle. For more detailed information (influence of pH, etc.) see Chapter X, § 2, p. 338.
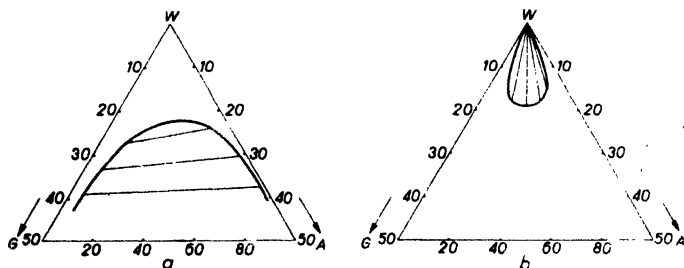
Next we must ask ourselves whether the difference in water-binding power of the two colloids brought forward by OSTWALD has in fact anything to do with hydration. We think this must be denied since although high colloid concentrations are necessary, yet in the 20—25% colloid mixtures so much water is still present that true hydration will probably have occurred completely.

The explanation of the partial miscibility phenomena will thus much rather have to be sought in the individually differing properties of the skein shaped kinds of macromolecule, which have an influence on the outcome of their struggle with each other to appropriate the water present in deficit as occlusion water.

We must however not forget that the macromolecules of gum arabic and gelatin are set along the molecule chain with numerous different groups which are polar (contain dipoles). For these kinds of molecules the considerations on the randomly kinked molecule in Chapter IV certainly do not hold rigidly, unless perhaps in extremely dilute sols. It is probable that the molecular skeins of these biocolloids differ from the theoretical ideal of Chapter IV by the additional statistical occur-

---

[1] There is agreement with this assumption in the success one has had in preparing totally closed gelatin bodies which enclose a vacuole containing gum arabic. The plan of the preparation was based on the behaviour on wetting of organic liquids immiscible with water by coacervates, see p. 437 Ch. XI § 1 e. H. G. BUNGENBERG DE JONG and O. BANK, *Protoplasma*, 33 (1940) 512.

rence of intra and intermolecular associations between these dipole groups, associations which will gain in importance as the molecular skeins have not sufficient space to unfold completely, that is to say, in concentrated sols.

It is a question in a mixture of such kinds of macromolecule which of the two has the greater tendency to such intra and intermolecular associations. Our conjecture that the gelatin-rich layer is present in this case as the coacervate, would indicate that the gelatin has the greater tendency in this direction than the gum arabic, which one is inclined to attribute to the tendency to associate of the peptide groups [1].

[1] Compare A. M. BUSWELL, W. H. RODEBUSH and M. F. ROY, *J. Amer. Chem. Soc.*, 60 (1938), 2444, who found a striking similarity of the infrared spectrum of mono substituted amides with that of gelatin. The association depends in both cases on a hydrogen bonding which is associated with enolisation of the —CONH— group.

IX. REVERSAL OF CHARGE PHENOMENA, EQUIVALENT WEIGHT AND SPECIFIC PROPERTIES OF THE IONISED GROUPS

H. G. Bungenberg de Jong

Department of Medical Chemistry, University of Leyden

§ 1. RECIPROCAL HEXOL NUMBER AND EQUIVALENT WEIGHT; CHARACTERISTIC CHARGE ELEMENTS

a. Introduction

Having studied in Chapter VII the charged kinetic units based on titrimetric, electrophoretic and viscosimetric data of sols, in this chapter further information regarding the ionised groups of colloids with electrolytic nature is obtained from reversal of charge phenomena occurring at *phase boundaries* or *film boundaries*. The method applied is the measurement of the electrophoretic velocities of small floccules or small coacervate drops, separated from the original sol by adding salts in increasing concentrations.

If however the salt employed does not cause flocculation or coacervation, the electrophoretic velocity is measured of small particles (e.g., SiO_2) suspended in the sol and covered by a complete colloid film (see p. 277 § 2 b).

The phenomena studied here are quite different from the well-known reversal of charge phenomena occurring in colloids of an ampholytic nature (e.g., proteins) on changing the pH. H ions will not interest us here, but all other kinds of cations (inorganic, organic and even macromolecular cations) acting at (nearly) constant pH on negatively charged colloids, further anions in general (OH ions excluded) acting at (nearly) constant pH on positively charged colloids.

The need for investigating these reversal of charge phenomena with ions other than H or OH, originated from a study of the so-called "complex colloid systems", dealt with in the next chapter (Ch. X). As the results obtained have a more general character, they have been brought together in this chapter.

Indeed the said reversal of charge phenomena reveal the existence of at least two charge elements, which play an important rôle in determining both the general and specific behaviour of colloids.

These two elements are the frequency of occurrence and the specific nature of the ionised groups on the macromolecule.

In the case of the corpuscular form of the macromolecule, supposing all ionised groups to be situated at the periphery of the corpuscle this frequency would have the significance of a *"density of charge"*, i.e., a *frequency per unit of surface*.

In the other extreme case, in which the macromolecule is not tightly folded up, but assumes the form of a randomly kinked long (eventually ramified) thread, this frequency is only that of a *frequency per unit length*.

The "corpuscular" type of macromolecule seems at first sight the simpler case,

as the relatively dense kinetic unit resembles a polyvalent ion, the charges being fixed at definite places. However in the known examples many complications occur, annihilating its seeming simplicity. For the corpuscular form of macromolecule is as yet only known to exist in native proteins, which are amphoteric substances. Thus positive and negative ionised groups occur here side by side, and only in media of extreme pH values will the kinetic unit have the character of a polyvalent cation or a polyvalent anion, this depending on the dissociation constants of acidic and basic groups. This will at least be the case at sufficiently low pH values, the COOH groups (which as a rule are the only acidic groups present) being then no longer dissociated. At sufficiently high pH values however in very many cases the kinetic units will not be strictly polyvalent anions because of the amino acid arginine almost always present as a constituent of the molecule. For at the end of the corresponding side chains of the macromolecule guanidine groups are then present, which group behaves as a very strong base. Thus the corresponding positively charged guanidinium groups are always present side by side with the negatively charged COO-groups.

A further complication in proteins is, that several groups of basic nature are very often present, not only arginine but also lysine and histidine frequently entering as constituents of the macromolecules, leaving respectively a NH_2 or an imidazol group free at the end of the corresponding side chains. In casein further phosphate groups are present side by side with the carboxyl groups.

Taking into consideration that all these different ionogenic groups each characterised by its specific properties, if situated at the surface of the corpuscular macromolecule, may occupy different relative positions, it will be evident that the corpuscular proteins cannot at all be considered as a simple case, suitable for starting a discussion on the significance of characteristic charge elements.

In this Chapter we shall therefore mainly confine ourselves to non corpuscular colloids, possessing only acidic groups, though proteins will also be considered shortly (after we have gained certain general view-points in these simpler cases).

Returning now to the frequency of occurrence of ionised groups on the macromolecule, the expression "density of charge" is not a particularly adequate expression, as this remind one of older points of view in Colloid Science, in which the kinetic units of the highly viscous type of colloids were symbolised as spheres, having a well-defined outer surface.

From the theory of the kinked macromolecule we have certainly to replace this by another expression, which defines the occurrence of ionised groups along the main chain of the high polymer. In Fig. 1 parts of (linear) macromolecules are represented in the streched form, each monomeric residue being symbolised by a rectangle, and each ionised group by a black dot.

It is assumed that the ionised groups, all of the same electrical sign, are distributed regularly along the main chain (which is of cause only a simplifying assumption). The frequency of occurrence of ionised groups could thus be expressed in A by $1/_\infty = 0$, in B by $1/_5$, in C by $1/_3$ in D by $1/_2$ and in E by 1, these figures representing the number of total ionised groups as a fraction of total number of monomeric residues.

It would be very interesting to investigate on such homogeneous series of polymers (the monomeric residues being the same and the ionised groups being of different types though within each series also the same) as regards the subjects, which form the contents of this Chapter.

The author and his coworkers in starting the study of reversal of charge pheno-mena used available biocolloids of acidic nature, in which thus different types of high polymeric molecules and different types of acidic groups (carboxyl, phosphate, or sulphate groups) are present.

As the structure is not or only aproximately[1] known for many of these bio-colloids it is not possible to compare the results with the above indicated rational frequency fractions for the ionised groups.

Fig. 1. Macromolecular colloids of acidic nature arranged in a series of increasing linear charge density (decreas-ing equivalent weight) considered at a pH, at which the ionogenic groups are completely ionised.
rectangles: □ monomeric residues.
black dots: ● ionised groups (negatively charged).
It is assumed for the sake of simplicity that each monomeric group can carry at most only one ionogenic group and that when there are fewer ionogenic groups than monomeric residues, the ionogenic groups are uniformly distri-buted along the chain molecule.

A: limiting case: ionogenic groups absent, that is to say, linear charge density $= 0$ and equivalent weight $= \infty$.
B: linear charge density and equivalent weight both finite.
The linear charge density increases by way of C and D to E (limiting case) to a maximum and the equivalent weight decreases to a minimum. If the ionisation is decreased by lowering the pH (for example $COO^{\cdot} \rightarrow COOH$), then one and the same colloid can pass through a displacement of its charge density or equivalent weight in the opposite sense (for example $E \rightarrow D \rightarrow C \rightarrow B \rightarrow A$ or $C \rightarrow B \rightarrow A$).
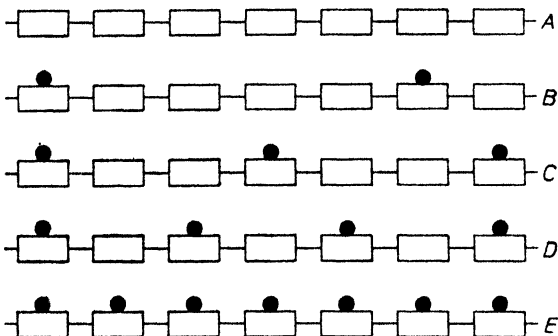
As a substitute we may use the frequency of the ionised groups per unit weight of the macromolecule, that is the reciprocal value of the equivalent weight, which may be determined analytically.

As will be shown in the following sections the value of the equivalent weight is already of the greatest importance in determining the differences in the behaviour of different colloids.

Apart from this first characteristic charge element, the equivalent weight, which determines the *number* of ionised groups per unit weight, we have to consider in the second place the *individual nature* of the ionogenic groups, the latter being especially important in explaining other specific differences in colloids of elec-trolytic nature.

Although those two characteristics also apply in principle to corpuscular native proteins, in this case a third characteristic element is present, viz., the fixed relative position of the ionised groups (the *individual pattern* of the charged groups) which may in special cases be far more important for the specific behaviour of the protein than the two first named characteristics. It is evident that, for example, the biological

---

[1] See schematic structure of these colloids as presented on p. 187-188, Ch. VII, § 1c.

specifity of proteins can only be understood on the basis of the very special geo-metrical arrangement.

As, however, data on the specific pattern are still extremely scarce and vague, this third characteristic element will not be considered in the sections to come.

Reversal of charge phenomena of association colloids have been found to follow the same rules as for "linear" macromolecular colloids. For this reason they are also discussed in this chapter and not in the chapter on association colloids (Ch. XIV), which is dedicated mainly to the association phenomenon itself.

By incorporating thus the reversal of charge phenomena of a series of phosphatide preparations a very welcome increase is obtained in the number of colloids, which derive their negative charge from the dissociation of phosphate groups.

b. The reversal of charge of sodium arabinate and other colloids of acidic nature with hexol nitrate. Reciprocal hexol number and equivalent weight

The biocolloids of acidic nature, the charge properties [1] of which will be discussed in this Chapter, are investigated in their usual form as salts. Thus for instance the natural gum arabic is mainly the Ca salt of a corresponding high molecular acid called arabinic acid; i⁺ contains also Mg, K and Na, though hardly any free elec-trolytic impurities are present. Its "ash content" is thus not the expression of impurity, but is the necessary results of the cations, bound to the ionised carboxyl groups present. By a simple procedure preparations can be obtained starting from the natural gum arabic, in which the original mixture of cations is replaced by one known kind only of cations, e.g., sodium arabinate, containing only Na-ions [2].

Such preparations can be analysed for their content of metal, which will enable one to calculate their equivalent weights. Sodium arabinate [2] gave in this way the mean value of 1202, a figure which after subtraction of 22 (i. e. — Na + H) is in good agreement with the equivalent weight of arabinic acid, viz., 1177, found by THOMAS and MURRAY [3] from titration curves of the latter substance in solution.

As was already stated in Ch. VII, p. 223, hexol nitrate, a complex cobalt salt with hexavalent cation, flocculates or coacervates most biocolloids of acidic nature [4], thus also gum arabic and sodium arabinate sols. Here the remarkable situation is encountered, that the hexol nitrate concentration needed just to start coacervation is nearly proportional to the arabinate concentration, this fact already indicating that the expression hexol nitrate "concentration" is here of a doubtful value: the hexol cations cannot be freely present in solution but must combine with the arabinate.

---

[1] H. G. BUNGENBERG DE JONG and F. H. TEUNISSEN, *Kolloid-Beihefte*, 47 (1938) 254. In an earlier investigation it had already been found, that different metal-arabinates give on analysis nearly the same figure for the equivalent weight of arabinic acid, this already showing that the cations present in the preparations are really involved in salt formation.

[2] Principle of the procedure: Repeated precipitation of a 10% solution containing also 10% NaCl with the two-fold volume of alcohol; thereafter once precipitation of the dissolved substance without adding salt, with a sufficient quantity of alcohol and at last repeated washing of the precipitate with ± 85% alcohol until NaCl is practically removed. H. G. BUNGENBERG DE JONG and P. VAN DER LINDE, *Bioch. Z.*, 262 (1933) 161.

[3] THOMAS and MURRAY, *J. physic. Chem.*, 32 (1928) 676.

[4] Exceptions glycogen and soluble starch (see Ch. VII, p. 208).

We have also already stated in Ch. VII, p. 220 that hexol nitrate brings about reversal of charge in all colloids of acidic nature. An indirect indication is the viscosity minimum obtained in soluble starch sols (see pag. 207, Fig. 19), but in sols of other colloids which are flocculated, the charge reversal can be followed by direct measurements of the electrophoretic velocity. At a hexol nitrate "concentration" just sufficient to start flocculation or coacervation, the floccules or coacervate drops are negatively charged, at higher concentrations they reach the reversal of charge point and become positively charged at still higher hexol nitrate concentrations.

Using graphs, in which the electrophoretic velocity is plotted as a function of the hexol nitrate concentration the reversal of charge concentration of the latter can be accurately determined, owing to the steep character of the curves.

If now dilute sodium arabinate sols of different concentrations are investigated, it appears that the hexol nitrate "concentration" required to reach exactly the reversal of charge point, is a strictly linear function of the colloid concentration, as shown in Fig. 2. The straight line obtained extrapolated to zero colloid concentration intersects the ordinate axis at very small positive value (4.10^{-6} N) thus indicating that we obtain two different kinds of information regarding the reversal of charge phenomenon.

a. the *real reversal of charge concentration*, i.e., the small concentration of free hexol

Fig. 2.   Reversal of charge of Na arabinate with hexol nitrate as a function of the sol concentration.
Ordinates: Gross reversal of charge concentrations in 10^{-4}N.
Abscissae: Arabinate concentrations in % by weight.
From the slope of the straight line one calculates that the reciprocal hexol number (number of grams of Na arabinate which binds one equivalent of hexol ions at the reversal of charge point) amounts to 1068. The true reversal of charge concentration (segment of the ordinate axis cut off by the straight line) amounts to only $0.4 . 10^{-6}$ N.

nitrate present in solution, indicated on the ordinate axis as the distance between the origin and the intersection point with the straight line.

b. the *bound quantity of hexol nitrate (or better of hexol cations) per gram sodium arabinate* at the reversal of charge point of the latter, which quantity can be calculated from the steepness of the straight line in Fig. 2.

Thus the hexol nitrate "concentration" needed for an arabinate sol of given concentration consists always of two parts viz., the real reversal of charge concentration and a fictitious concentration — the quantity fixed by and proportional to the arabinate present.

At moderate arabinate concentrations (right side of the figure) the hexol nitrate "concentration" is thus for the greatest part fictitious, only at very small arabinate concentrations, (to the extreme left in Fig. 2) it would be practically equal to the real reversal of charge concentration.

Of these two pieces of information the former (a) does not yet concern us here, it will however play a great rôle in considering reversal of charge phenomena with lower valent cations or of positively charged colloids with lower valent anions in the §§ 2 - 6 of this Chapter.

Information (b) can be defined as the "Hexol number" of Na-arabinate. However its reciprocal value, the so-called "*Reciprocal hexol number*" is of more value for us. The reciprocal hexol number of a colloid is thus defined as the quantity in g of the colloid for which there is needed exactly one gram equivalent of bound hexol cations to reach the reversal of charge point.

From the data represented in Fig. 2 the reciprocal hexol number of Na-arabinate was calculated to be 1068. Comparing this with the analytically found equivalent weight (see above), for which 1202 was obtained, it appears that both values are of the same order, but are not equal, the reciprocal hexol number being smaller than the equivalent weight.

The results obtained with Na-arabinate do not stand apart, but a similar behaviour as regards the reversal of charge with hexol nitrate is obtained with all other investigated colloids of acidic nature. As in Fig. 2, in all cases straight lines are obtained, intersecting the o᷎ ᷎ ᷎᷎᷎᷎ axis in most cases at very small valu᷎᷎, not exceeding a few units of $10^{-5}N$. The straight lines have however different slopes, see for instance Fig. 3, indicating that the reciprocal hexol number varies considerably from one colloid to another.

Table 1 shows that in the colloids investigated the equivalent weight increases twelvefold and that the reciprocal hexol number also increases to the same order. But as with Na-arabinate the latter is always smaller than the former. Thus apart from a systematic difference — which is discussed in the next subsection — the

Fig. 3. Reversal of charge of Na nucleate with hexol nitrate as a function of the sol concentration.
Ordinates: Gross reversal of charge concentrations in 10^{-4} N.
Abscissae: Nucleate concentration in % by weight.
From the slope of the straight line one calculates that the reciprocal hexol number (number of grams which binds one equivalent of hexol ions at the reversal of charge point) amounts to 294. The true reversal of charge concentration (segment of the ordinate axis cut off by the straight line) amounts to only $0.6 \cdot 10^{-5}$ N.

reciprocal hexol number has the significance of an equivalent weight determined by direct electrophoretic measurements. So this method could be used in general to determine approximately the equivalent weight of colloids for which a direct analysis has not yet been made, or is technically difficult. Then the reciprocal hexol number is determined and the value obtained multiplied by 100/84.

TABLE 1

Substance	Reciprocal hexol number	Equivalent weight	Reciprocal hexol number / Equivalent weight	
Na-pectate	203	233	0.87	
Na-carrageen	233	294	0.79	
K-chondroitin sulphate . .	290	319	0.91	
Na-yeast nucleiate	294	373	0.79	mean
Soya bean phosphatide II* .	782	988	0.79	0.84
Na-pectinate	1040	1157	0.90	
Na-arabinate	1068	1202	0.89	
Na-agar	2264	2909	0.78	

In the above table the substance marked with asterisk — an alcohol insoluble fraction of the total soya bean phosphatides — has been included.

In column 3 this association colloid has been characterized by its apparent equivalent weight (p. 273, § 1 f) and it then shows quite the same behaviour as the other substances which are macromolecular colloids (p. 188).

Very often association colloids will show a quite similar behaviour to macromolecular colloids as regards other charge properties too.

As therefore no real differences between the two kinds of colloids exist on this point, association colloids will also be frequently considered in the following pages, in as much as this makes a survey of general principles easier.

Instead of the hexavalent hexol ion lower polyvalent cations could also be used in principle

Fig. 4. Reversal of charge of alcohol soluble soya bean phosphatide with hexol nitrate or with La(NO₃)₃, both as a function of the sol concentration.
Ordinates: Gross reversal of charge concentration in m. eq. per l.
Abscissae: Phosphatide concentration expressed in that of the initial sol (= 0.858%) as unit.
The two straight lines are practically parallel, that is to say, for the charge reversal mutually equivalent amounts of La and hexol ions are bound per gram of phosphatide. The true reversal of charge concentrations are given by the segments which the straight lines cut off from the ordinate axis. This is very much greater for La(NO₃)₃ than for hexol nitrate.
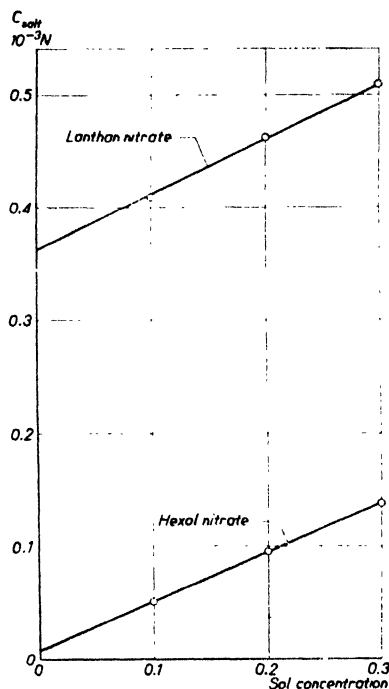
for determining an analogous reciprocal cation number. From a practical point of view however this substitution offers only disadvantages. See for instance Fig. 4 in which a sol of a certain phosphatide preparation (an alcohol-soluble fraction of the total soya bean phosphatides) has been investigated both with hexol nitrate and La(NO₃)₃ [1].

---

[1] H. G. BUNGENBERG DE JONG, H. L. BOOIJ and J. G. WAKKIE, *Kolloid-Beihefte*, 44 (1936) 254.

The figure shows the general effect in choosing a lower valent cation, which consists in a nearly parallel displacement of the straight line towards higher "concentrations." The nearly equal slope of both lines means that the reciprocal La number will indeed be practically the same as the reciprocal hexol number. The main effect consists thus in an enormous increase of the real reversal of charge concentration, the latter being 7×10^{-6}N for hexol nitrate and $3.6 \cdot 10^{-4}$N for La(NO$_3$)$_3$.

This increase however means a great disadvantage interfering considerably with the accuracy of the calculated reciprocal La number.

For this number must now be obtained from relatively not very different gross reversal of charge concentrations, which latter should for this purpose be known with much greater accuracy than in the case of hexol nitrate. Now in general choosing a lower valent cation, the slope of the electrophoretic velocity-concentration curve decreases, so that in fact the gross reversal of charge concentration with La(NO$_3$)$_3$ can be determined with less accuracy than with hexol nitrate.

Thus the accuracy of a calculated reciprocal La number must be necessarily much lower than of a reciprocal hexol number.

Now we have taken as an example a still very favourable case, phosphatides, showing relatively low gross reversal of charge concentrations with La(NO$_3$)$_3$. In other colloids the latter are as a rule much higher, so that the applicability of La for obtaining a reciprocal La number is experimentally quite excluded.

One should thus choose a higher valent cation. But the tetravalent Th ion is not very suitable, for complications may occur as a result of considerable hydrolysis in the small concentrations in general used in this kind of experiment.

It is for avoiding hydrolysis, that in many experiments polyvalent complex cations are to be preferred in place of polyvalent monoatomic cations.

Summarizing we may say that hexol nitrate is at the moment the salt best suited for the experimental method discussed in this subsection, the hexol cation by its high valency giving very low real reversal of charge concentrations and by its nature of a strong complex ion being hardly hydrolysed at low concentration (see also p. 300, note 1).

Still hexol nitrate has its inconveniences, its solutions decomposing as time proceeds. Therefore hexol nitrate solutions should always be used perfectly fresh.

c. Origin of the negative charge of colloids of acidic nature and mechanism of the reversal of charge with hexol nitrate.

The results discussed in § 1 b give strong evidence that the electric charge of colloids of acidic nature takes its origin from the electrolytic dissociation of the ionogenic groups present.

Still the systematic difference between the numerical values of reciprocal hexol numbers and equivalent weights calls for an explanation. This discrepancy is not easily explained from considerations in homogeneous media.

Then it would be expected, that an amount of fixed hexol ions, which is exactly equivalent to the amount of ionised groups present, will suffice in annihilating the negative charge of the colloid, for instance, of the arabinate. From the experimental fact that the reciprocal hexol number is smaller than the equivalent weight, it follows however that for reaching the reversal of charge point a quantity of hexol cations must be bound on the arabinate which is somewhat greater than the equivalent amount.

For an explanation[1] we must start from the fact, that the hexol arabinate formed does not remain in homogeneous solution but separates, and that the reversal of charge measured relates to the electrical proporties of the phase boundary of the separated hexol arabinate.

---

[1] H. G. BUNGENBERG DE JONG and P. H. TEUNISSEN, *Kolloid-Beihefte*, 47 (1938) 254.

Further it must be stated here, that though this separation of hexol arabinate resembles a double decomposition:

Na-arabinate + hexol nitrate → hexol arabinate + $NaNO_3$,
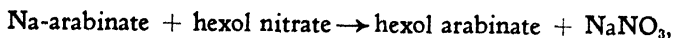
this formulation is still too simple. The separated phase (see p. 385 Fig. 35) is in no sense a solid of definite composition, but a coacervate i.e., a colloid-rich liquid (for fuller particulars on coacervates see Chapter VIII). Apart from a considerable (and variable) amount of water, the coacervate has a variable composition as to its essential components, here arabinate ions and hexol ions[1]. Dependent on the relative amounts of Na arabinate and hexol nitrate present in the total system, the coacervate will contain, apart from hexol arabinate proper, an excess of Na arabinate or of hexol nitrate, see formulations below, A and C). Thus only at a quite definite relative proportion of Na arabinate and hexol nitrate in the total system the coacervate will contain equivalent amounts of hexol ions and arabinate ions (formulation B).

A. (excess of Na arabinate)   ⌈ hexol arabinate ⌉   ⌒
                               | Na arabinate | ⌒ arabinate`   Na`.
                               ⌊ p H_2O ⌋   ⌣

B. | hexol arabinate |
                               ⌊ q H_2O ⌋

C. (excess of hexol nitrate) ⌈ hexol arabinate ⌉   ⌒
                               | hexol nitrate | ⌒ hexol`      $NO_3$`.
                               ⌊ r H_2O ⌋   ⌣
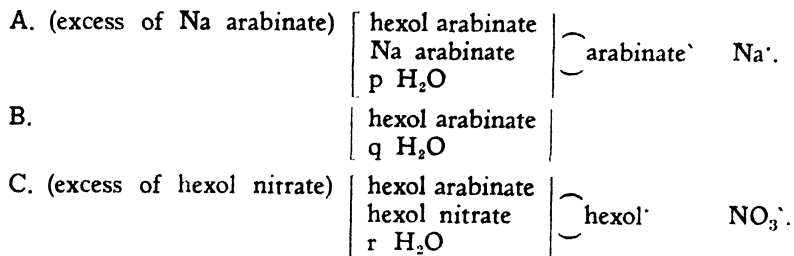
Now returning to the method used in determining the reversal of charge point, measuring the electrophoretic velocity of the minute separated coacervate drops, it will be clear that we must particularly take into consideration the electrical properties of the surface of the coacervate drops.

That the latter appear negatively charged if not enough hexol nitrate is added, that is so long as Na arabinate is present in excess, indicates that at the surface of the coacervate drops Na arabinate is adsorbed in the way indicated in A[2].

The positive charge of the droplets is in the same way caused by adsorption of hexol nitrate at the surface of the coacervate droplets as indicated in C[2].

We come now to the difficulty proper, mentioned already above viz., that the reciprocal hexol number is somewhat smaller than the equivalent weight.

If we assume that the coacervate which contains only hexol arabinate has no double layer at the surface of its drops as indicated in B, then reciprocal hexol number and equivalent weight should be exactly equal.

For explaining that the former is smaller than the latter, we must assume that this coacervate of equivalent composition still carries a negative charge at the surface of the coacervate drops.

---

[1] This has not been determined analytically, but has been deduced from the close analogy between coacervates of the type colloid anion + colloid cation and coacervates of the type colloid anion + crystalloid cation (p. 384 Ch. X § 3). For the variable composition of the former type see p. 355-364 Ch. X.

[2] In the above symbols the arabinate and hexol ions are written as monovalent ions for simplicity. Of cause their polyvalency will essentialy contribute to their adsorption at the surface of the coacervate as indicated in the symbols.

If hexol arabinate situated at the surface of the coacervate drop has the tendency to dissociate hexol ions such a negative charge of the surface can indeed arise. We thus arrive at the following formulation D for the coacervate of equivalent composition:

D. (equivalent coacervate)   ⌈ hexol arabinate ⌉⌒arabinate`  hexol`
                               ⌊ q H₂O   ⌋⌣
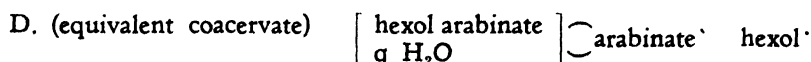
Of course such an assumption must also be made in the cases A and B, the actual charge on the surface being the result of a negative charge as in D, together with the negative or positive charge resulting from an excess of Na arabinate or hexol nitrate in the total system as in A and C.

The reversal of charge point can then be formulated by E, in which the negative charge of D is just compensated by the positive charge of hexol nitrate adsorbed at the surface.

E. (reversal of charge point) ⌈ hexol arabinate ⌉⌒arabinate : hexol`
                               | hexol nitrate |⌒
                               ⌊ q` H₂O          ⌋⌒hexol`   NO₃`
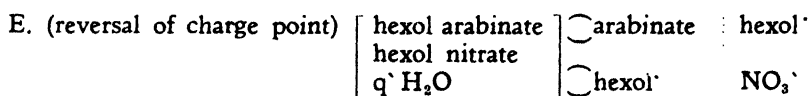
This extra adsorbed hexol nitrate on the coacervate surface corresponds however to a coacervate which no longer exclusively contains hexol arabinate but also a certain amount of hexol nitrate (analogous to C). Thus at the reversal of charge point (E, or simplified to symbol F) the proportion hexol: arabinate is higher than the equivalent one (D).

F. (reversal of charge point) ⌈ hexol arabinate ⌉
                               | hexol nitrate |
                               ⌊ q` H₂O   ⌋
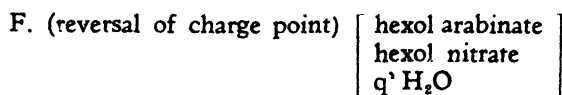
But the latter statement is identical with the experimental fact that the reciprocal hexol number is somewhat smaller than the equivalent weight.

In the above the systematic difference between reciprocal hexol number and equivalent weight was discussed for an example in which the "separated phase" had the nature of a coacervate. As the explanation is founded on certain general phenomena at phase boundaries, the considerations hold also for coacervates which contain more water or less water than in the example, still for coacervates poor or very poor in water, which morphologically appear as granular or really flocculent precipitates — and even for flocculations in which the floccules have certainly not a coacervate nature — as for carrageen, where the floccules appear under the microscope as stringy or fibrillar masses. Further it may be added, that the discussion does not contain specific elements which would only apply to hexol ions. Thus in general it may be expected, that reversal of charge of negative colloids by other cations will also occur at higher concentrations than those needed to attach an equivalent amount of cations to the ionised groups.

This may possibly explain the fact that the viscosity minima in Fig. 33 (see p. 221) lie at lower concentrations than the reversal of charge concentrations.

In view of considerations on "Complex Systems" to be discussed later (Ch. X) it seems plausible to assume that at the viscosity minima the arabinate has fixed an equivalent amount of cations.

The reversal of charge concentrations were however determined on the electrophoretic velocity of suspended SiO_2 particles, because the salts used did not flocculate the arabinate. The SiO_2 particles carry an adsorbed arabinate film (see p. 277, § 2 b). Nevertheless the above considerations on phenomena at the boundary of a colloid rich phase must also apply for the boundary of an adsorbed colloid film on SiO_2 particles. If the composition of the adsorbed film corresponds to an equivalent amount of arabinate and fixed cations, the film will still carry a negative charge by the dissociation tendency as in D. The reversal of charge will thus be found at a higher concentration than the viscosity minimum.

Reversal of charge phenomena are not restricted to combinations of the type, (negative) colloids of acidic nature + cations. They are also found in amphoteric colloids, for instance negatively charged proteins with cations and positively charged proteins with anions. Here also by choosing an appropriate polyvalent cation (hexavalent hexol) or anion (tetravalent ferrocyanide or hexavalent Germanin[1]) corresponding reciprocal cation numbers or reciprocal anion numbers may be obtained. They may further be compared with their "equivalent weights" — these being, in principle, of an algebraic character, as positively and negatively charged ionised groups may be present side by side (p. 273, § 1 f). Examples (casein at pH 7.6, 7.0, 6.7, 6.1, 3.4, and 2.9; clupein at pH 3.4) have been investigated by L. TEUNISSEN - VAN ZYP[2].

The systematic differences between reciprocal ion numbers and equivalent weights have been found here also to be of the same order of magnitude as those in Table 1 (p. 265), the former being 10—15% smaller than the latter.

d. Equivalent weight and flocculability.

In studying flocculation viz., coacervation phenomena of biocolloid sols occurring with added salts, the author and coworkers have found that relatively rarely they have the character of salting out phenomena (e.g. agar with Na_2SO_4 or $MgSO_4$ in high concentrations), but in most cases they show striking analogies with the mutual flocculation, viz., coacervation phenomena of oppositely charged colloids (for fuller information see Chapter X).

In salting out agar with neutral salts the lyotropic action of both salt-ions, especially of the anion is of primary importance. Further the salting out occurs only at a rather high concentration.

In contrast to this the coacervation of Na arabinate with hexol nitrate occurs already at very low concentrations and is the consequence of the mutual electric attraction between the negative arabinate ion and the polyvalent hexol cation.

We have already seen in Ch. VII § 10 p. 223) that, as regards the flocculability of biocolloids of acidic nature with salts, distinct specific factors play an important rôle.

Now it appears that the equivalent weight must be regarded as the most conspicuous factor in determining the specific behaviour of biocolloids, this being a reciprocal measure of the frequency of occurrence of the ionised groups[3].

---

[1] Germanin, or Bayer 205, the sodium salt of a complicated derivative of urea, containing six SO_3H groups.
[2] L. TEUNISSEN-VAN ZIJP, Thesis, Leiden 1938.
[3] H. G. BUNGENBERG DE JONG and P. H. TEUNISSEN, Kolloid-Beihefte, 47 (1938) 254.

This is shown in the following Table 2, in which fourteen colloids are listed in the order of decreasing reciprocal hexol number. The table contains information on the flocculability with six different salts, viz., hexol nitrate (6—1); rhodochromium chloride (5—1): platinium triethylenediamine tetranitrate (4—1); luteocobalt chloride (3—1); $CaCl_2$(2—1) and NaCl(1—1).

In the table — means no flocculation or coacervation; + opalescent systems; + + distinct flocculation or coacervation.

TABLE 2

COLLOID	R. H. N. (Reciprocal hexol number)	6—1	5—1	4—1	3—1	2—1	1—1	$\dfrac{10^3}{R.H.N.}$
Glycogen [1]	78 000	—	—	—	—	—	—	0.013
Soluble starch [1]	26 000	—	—	—	—	—	—	0.039
Egg lecithin [1]	20 000	+ ??	—	—	—	—	—	0.05
Soya bean phosphatide I	4 327	+ +	+ +	+ +	—	—	—	0.248
Na-agar	2 264	+	+	+	—	—	—	0.442
Na-arabinate	1 068	+ +	+ +	+ +	—	—	—	0.936
Na-pectinate	1 040	+ +	+ +	+ +	—	—	—	0.962
Soya bean phosphatide II	782	+ +	+ +	+ +	+ +	+ +	—	1.28
Na-semen lini mucilage	563	+ +	+ +	+ +	+ +	—	—	1.78
Na-thymus nucleiate	326	+ +	+ +	+ +	+ +	+ +	—	3.07
Na-yeast nucleiate	294	+ +	+ +	+ +	+ +	+ +	—	3.40
K-chondroitin sulphate	290	+ +	+ +	+ +	+ +	—	—	3.45
Na-carrageen	233	+ +	+ +	+ +	+ +	—	—	4.29
Na-pectate	203	+ +	+ +	+ +	+ +	+ +	—	4.93

Inspecting the above table, we see that the first three colloids with very high reciprocal hexol numbers do not flocculate with either of the six salts. The next colloids with reciprocal hexol numbers ranging from 5 000—1 000, do flocculate or coacervate with 6—1, 5—1 and 4—1.

The colloids with very low reciprocal hexol numbers show flocculation or coacervation with 6—1, 5—1, 4—1, 3—1 and in some instance even with 2—1. Thus it seems that the equivalent weight (which multiplied by approximately 0.85 is the reciprocal hexol number) is an important factor in determining flocculability with salts of the types considered here.

With decreasing equivalent weight of the colloid flocculability increases, in so far as flocculation or coacervation is already realisable with lower valent cations the lower the equivalent weight (see however also p. 295, § 2 m).

The table contains three phosphatides, which do not belong to macromolecular but to association colloids. For including them here see p. 262 and 265.

---

[1] As no flocculation, occurs the reversal of charge concentrations needed for the calculation of R.H.N. were determined on suspended particles (for instance SiO_2) which become covered with a complete colloidal film. For particulars see p. 277 § 2 b.

e. Indications of the occurrence of a second characteristic charge element, determining specific behaviour of colloids

On closer inspection of Table 2 on page 270, we see that the correlation between reciprocal hexol number and flocculability is not a rigorous one, so, for instance, Na agar though of lower reciprocal hexol number than the preceding Soya bean phosphatide, shows only opalescence with 6—1. 5—1 and 4—1.

Further irregularities occur in the lower part of the table as regards the flocculability with 2—1, which should be met with only in colloids having the lowest reciprocal hexol numbers.

In Fig. 5 (upper) the data of the table are represented graphically, each colloid being represented by a point of which the ordinate and absciss need further explanation.

The ordinate value is that salt in the series 6—1, 5—1, 4—1, 3—1, 2—1, 1—1, which, going from left to right in this series, is the first which does not flocculate or coacervate the biocolloid considered.

Thus soluble starch is characterised by 6—1, Na arabinate by 3—1 and Na pectate by 1—1.

As abscissa values are taken not the reciprocal hexol numbers (R.H.N.) themselves, but the thousand-fold values of their reciprocals, thus $\dfrac{10^3}{\text{R.H.N.}}$.

This value can be considered as an approximate measure of the charge "density" (or in randomly kinked macromolecular colloids as a measure of frequency of occurrence of the ionised groups, see p. 259 § 1 a).

The different points so obtained, each characterising a biocolloid as to its "flocculability" and "charge density", can be united to give *three different curves* as in Fig. 5 (upper).
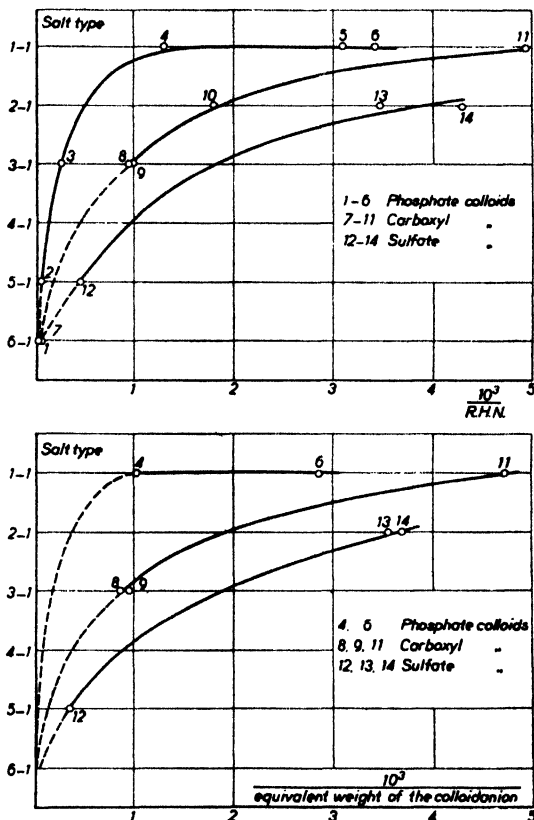
Fig. 5. Relation between tendency to flocculation or coacervation and reciprocal hexol number (upper) or equivalent weight of the colloid anion (lower). (See text).

1. soluble starch (Merck)
2. egg lecithin
3. soya bean phosphatide
    I (alcohol soluble)
4. soya bean phosphatide
    II (alcohol insoluble)
5. Na thymus nucleate
6. Na yeast nucleate
7. glycogen (Schuchardt)
8. Na arabinate
9. Na pectinate
10. Na semen lini mucilage
11. Na pectate
12. Na agar
13. K chondroitin sulphate
14. Na carrageen

In doing so, the colloids are divided into three classes. Now it is interesting, that each of these classes comprises colloids with the same composition of the ionogenic groups.

On the upper curve are situated egg lecithin, both soya bean phosphatides thymus and yeast nucleates all having *ester phosphate groups*. They may be called *phosphate colloids*.

On the middle curve are situated Na arabinate, Na pectinate, Na semen linimucilage and Na pectate all having *carboxyl groups*. They will be called *carboxyl colloids*.

On the lower curve we find Na-agar, K-chondroitin sulphate and Na-carrageen all having *ester sulphate groups*. We will call them *sulphate colloids*.

We may conclude that flocculability is not only determined by the "charge density" *but in addition by the composition of the ionised groups.* At the same "charge density" the flocculability increases in the order:

$$\text{sulphate colloid} \ < \ \text{carboxyl colloid} \ < \ \text{phosphate colloid.}$$

It must be mentioned expressly, that these statements are based on the behaviour of colloids towards six randomly chosen salts. They have therefore no absolute validity as to flocculation with other salts expressed by the same cation-anion symbols (see p. 295, § 2 m). But they give us a first indication that for the specific behaviour of colloids a second characteristic charge element is involved, which is connected with the *composition of the ionised groups*.

In Fig. 5 (upper) glycogen and soluble starch have received the symbol 6—1 as "flocculability". Strictly speaking the symbol could perhaps also be 7—1 or higher, as salts of this type are not available to check this. As to the question to which types of colloids these two belong no definite answer can be given. According to SAMEC[1] native starch contains ester phosphate groups which are hydrolysed easily. In our preparation of soluble starch (MERCK) this hydrolysis may already have taken place during its manufacture, and a few carboxyl groups may have been formed by oxidation (an indication for this was the order of cations by precipitation with alcohol, which is that to be expected for a carboxyl colloid, see p. 400). Therefore the middle curve for carboxyl colloids has been drawn to the left up to the point characterising our soluble starch preparation. As Na agar gave no flocculation with 6—1, 5—1 and 4—1, but only opalescence the choice of a flocculability symbol is difficult here. We chose 5—1.

Egg lecithin sol gave with 6—1 a slight turbidity only very slowly and only in the neighbourhood of the reversal of charge point, therefore in the Fig. symbol 5—1 has been chosen for its flocculability.

As all three curves descend steeply in the left part of Fig. 5 (upper) the uncertainties in choosing flocculability symbols for agar, egg lecithin, glycogen and amylum, have no influence on the general aspect of this figure.

Fig. 5 (upper) has been drawn using as a measure of "charge density" the values $10^3/\text{R.H.N.}$

This choice has the advantage that all colloids listed in Table II, can be used. As the equivalent weight in macromolecular colloids is approximately $0.85 \times \text{R.H.N.}$, an analogous figure in which $10^3/\text{equivalent weight}$ is used as a measure of the charge density, cannot and does not differ materially in general aspect from Fig. 5. However such a graph (which is not reproduced here) comprises less points representing

---

[1] M. SAMEC, *Kolloidchem. Beih.*, 6 (1914) 23.

individual colloids, for analyses from which the equivalent weight could be deduced have not been made for the preparations no 1, 5, 7 and 13 of Fig. 5. Nor is this the case for the phosphatide preparations no 2 and 3 (where the apparent equivalent weight, see § 2 f, which is only of importance here could have been calculated from accurate analytic data, see p. 295).

The remaining eight colloids all belong to the macromolecular type. Considering the question why the equivalent weight might be of basic importance in colloid behaviour, one is rapidly led to the idea, that strictly speaking not the equivalent weight of the total colloid (this being a Na or K salt) itself, but the equivalent weight of the colloid anion must be of basic importance. Therefore Fig. 5 (lower) is drawn, in which instead of 10^3/equivalent weight an analogous expression is used, in which the equivalent weight is replaced by the value obtained by subtracting the weight of the cation present (Na or K) from the equivalent weight. Comparing Fig. 5 (lower) with Fig. 5 (upper) we notice, apart from a smaller number of points representing individual colloids (explained above), a striking resemblance, once more three different curves appearing, each characterised by colloids carrying phosphate, carboxyl and sulphate groups. From this Fig. 5 we may thus conclude that specific behaviour of macro- molecular colloids of an electrolytic nature is connected with two characteristic charge elements; viz., *equivalent weight of the colloid anion* (an approximate reciprocal measure of frequency of occurrence of the ionized groups along the chain molecule) and *composition of the ionised groups.*

It is most remarkable that Na yeast nucleate and Na thymus nucleate, colloids of very related structure but at the same time of very greatly different chain length of the macromolecule, (see p. 188 and 227), show so little difference in the "flocculation behaviour" expressed in the Figures 5 (upper) and 5 (lower), that their points representing them in Fig. 5, lie close together. This seems to underline that for this behaviour *not* the weight of the macromolecular anion as a whole is very important but just the above named frequency of occurrence of the ionised groups along the chain molecule.

Also chondroitine sulphate and mucoitine sulphate usually considered as relatively small molecules, so that our definition of these substances as colloids (Chapter I, p. 3) might be in danger, lay on the same curve characteristic of sulphate colloids as the truly long chain molecules of Na-agar and Na-carrageen.

f. Apparent equivalent weights

We may speak of an *apparent equivalent weight*, playing for colloid behaviour towards added salts a rôle similar to the true equivalent weight, if not all ionogenic groups are ionised but only a fraction of them.

The apparent equivalent weight is necessarily higher than the true equivalent weight, the frequence of occurrence of ionised groups along the chain molecule becoming smaller.

With a colloid of acidic nature a change in the indicated direction ("charge density" decreasing) takes place on sufficiently lowering the pH (see Fig. 1, p. 261).

In colloids of ampholytic nature, as proteins, the electrophoretic method for determining charge "density" (hexol nitrate in negative protein sols, germanin in positive protein sols) gives in many cases only *apparent equivalent weights*, because positive ionised groups (e.g., side chains of arginin, lysin, histidin) and negative

ionised groups (e.g., side chains of aspartic, glutamic, hydroxyglutamic acid and possibly phosphoric acid) are simultaneously present, one of these kinds preponderating over the other according to the position of the chosen pH with respect to the *I.E.P.*

The said apparent equivalent weights play a great part in complex coacervation of oppositely charged colloids, and explains for instance the shift in optimal mixing ratios by altering the pH (see p. 322, § 6 b and p. 359, chapter X, § 2 i).

Further the magnitude of the apparent equivalent weights of both colloids is here of great importance, the interaction being the more intense, the lower the apparent equivalent weight (see p. 374).

Proteins taking part in the formation of tricomplex colloid systems (see p. 415), act however preferably as amphoions. Therefore the above no longer applies here, for at the most favourable pH, namely the *I.E.P.*, the apparent equivalent weight is to be considered as infinite.

Phosphatide preparations, if present as non-sensitised[1] sols, show according to Table 2 and Fig. 5 (upper), a behaviour similar to macromolecular colloids, the experimental points fitting in remarkably well in the curve for phosphate colloids, if these phosphatide preparations are characterised by their reciprocal hexol numbers. We can thus attribute to such an phosphatide preparation (being a mixture of true phosphatide molecules or better phosphatide amphoions and a varying amount of phosphatidic anions) an "equivalent weight" ($\pm 0.85 \times$ reciprocal hexol number). It will be clear, that again we have to do with an *apparent equivalent weight*, this being the weight in grams of the phosphatide preparation, that contains one gram equivalent of ionised phosphate groups in excess. The mixture of the constituent parts (amphoions + phosphatidic anions) can here behave as a whole by virtue of strong association forces.

This apparent equivalent weight plays a great part a) in determining the spread of cations in the reversal of charge spectrum (§ 21, p. 295); b) in the extent of the antagonism $CaCl_2 - NaCl$ (§ 5 b, p. 314-315); c) in complex flocculation ‹ ‹‹mplex coacervation with positive protein sols (see p. 374 Ch. X § 2 r).

In tricomplex systems (Ch. X § 6 p. 415) however this no longer holds. Here the amphoions of the phosphatide preparation form the reacting component of the mixture, and the phosphatidic anions present are only a nuisance. Therefore the said tricomplex systems are the more typical the higher the reciprocal hexol number, that is the fewer phosphatidic anions are present.

g. Characteristic properties of the ionised groups causing specific behaviour of colloids

Having seen in § 1 e. that not only charge density, but also the composition of the ionised groups plays a rôle in determining specific behaviour of colloids, we will

---

[1] Different organic substances, (*e.g.*, triolein, cholesterol) if present in a phosphatide sol, bring about sensitisation, that is, the sol now flocculates or coacervates rapidly after addition of salts. By this sensitisation the above resemblance with macromolecular sols is lost, the sol already flocculating or coacervating with lower valent cations than would be expected from the value of the R.H.N. This shifting towards lower valent cations is the greater as the degree of sensibilisation is increased, so that eventually all cations, the monovalent ones included, will flocculate or coacervate. For example see: H. G. BUNGENBERG DE JONG and R. F. WESTERKAMP, *Biochem.-Z.* 248 (1932) 309.

now discuss which properties of the iniosed groups are responsible for it. Up till now two such properties have been recognised, viz.,

a. the different affinities of the ionised groups for H ions, characterised by the different dissociation constants of the ionogenic groups.

b. the different polarisabilities of the ionised groups, playing an important rôle in the fixation of inorganic ions on the ionised groups.

The nature of the ionogenic groups will of course have a great importance for the pH range in which electrolytic dissociation is possible.

Thus *"Carboxyl colloids"*, possessing only carboxyl groups, — the latter having a rather low dissociation constant — will lose their negative charge relatively easily by lowering the pH.

"Phosphate colloids" however will loose their charge only at a much lower pH. Finally the *"Sulphate colloids"*, possessing ester sulphate groups; loose their negative charge at a still lower pH, the sulphate groups having a still stronger acid nature than the phosphate groups.

At pH values not far from neutrality (pH 5—8) these different values of the dissociation *constants no longer play a rôle*, ionisation being practically complete even in the case of carboxyl colloids[1].

In accordance with it preliminary experiments[2] on the influence of pH on the reciprocal hexol number of arabinates showed that the value of the latter is practically constant in the pH range 6.2—4.5 and that only at lower pH values a change takes place in the expected direction.

In the next section, we shall discuss at some length the reversal of charge of colloids of acidic nature with a greater number of inorganic cations in neutral or slightly acid media, the ionogenic groups of all three kinds of colloids (phosphate, carboxyl, sulphate colloids) thus being practically wholly ionised.

We shall see that for these reversal of charge concentrations the simple rules no longer hold, which we met in studying the suppression of the electroviscous effect at very low salt concentrations. (see p. 203 Ch. VII § 5 b) Indeed the valency of the cation is no longer the only factor of primary importance, as very marked specific differences occur between cations of the same valency. Other properties of the cations viz., volume and polarising action come into play next to valency.

---

[1] The carboxyl group in the carboxyl colloids derived from polymer carbohydrates is distinctly stronger than that in fatty acids, due to the precence of many OH groups in the molecule. It has been mentioned in Ch. VII, § 2 b, p. 189, that the pH value is found to vary with degree of neutralisation (or pH), and salt concentration. The lowest values of pK found for pectic acid and gum arabic are of the order of 2.7. The extrapolated value for low charge or high concentration of salt is found by OVERBEEK (loc. cit. p. 189) to be about 2 to 2.5, whereas for a fatty acid the pK is of the order of 4.7. Consequently the dissocation constants differ by more than a factor 100.

[2] H. G. BUNGENBERG DE JONG and P. VAN DER LINDE, *Biochem. Z.*, 262, (1933) 162, in which publication an explanation for the systematic difference between reciprocal hexol number and equivalent weight was given, that cannot be uphold and is to be replaced by that given above in § 1 c. The observed change in reciprocal hexol number at pH values below 4.5 must therefore be explained as a decrease of the ionisation, as suggested above. The sample of hexol nitrate used, has later been shown to be not sufficiently pure, from which much too low reciprocal hexol numbers resulted. However the conclusion drawn that ionisation is practically complete in the pH range 6.2 — 4.5 seems not to be endangered by it.

Now it is interesting that the order in which cations follow one another as regards the magnitude of the reversal of charge concentrations is not always the same, three types occurring each characteristic of phosphate, of carboxyl and of sulphate colloids,

These different types, in which thus the composition of the ionised groups expresses itself, can be explained by assuming that the *polarisabilities* of the corresponding ionised groups and of water itself decrease in the following order:

Phosphate group $\rangle$ carboxyl group $\rangle$ H_2O $\rangle$ sulphate group.

For a further discussion, which seems only fruitful after expounding the experimental results more in detail, the reader should consult the next section.

§ 2. SPECIFIC CATION SEQUENCES IN THE REVERSAL OF CHARGE OF COLLOIDS OF ACIDIC NATURE WITH INORGANIC CATIONS

a. Introduction

It is the aim of this section to compare for a number of colloids the relative positions of the reversal of charge concentrations of a large number of inorganic ions.

Before giving experimental results and explaining them, it is necessary to make clear, that the kind of useful information thus obtained is quite different from that, which we obtained in § 1 from determinations of the reversal of charge with hexol nitrate.

We must remember, that the experimentally determined reversal of charge concentration, C, is strictly speaking only a gross concentration see p. 263, composed of two parts:

a. the true reversal of charge concentration, C_t *i.e.*, the concentration of the cation considered, present free in the medium, at the reversal of charge point.

b. a fictitious concentration, C_f, representing the amount of cations bound to the colloid present, at the reversal of charge point.

We may thus write very generally $C = C_t + C_f$, and it will thus depend on the relatively magnitudes of C_t and C_f, what kind of information the reversal of charge concentration mainly gives.

The general equation may for instance approximate to $C = C_f$, which at moderate sol concentrations is the case for hexol nitrate, C_t here being very small. Therefore hexol nitrate is a salt which is well fitted to give information on the "charge density" of colloids as was discussed in details in § 1. In that section C_f was the only quantity which really interested us, leading to the calculation of reciprocal hexol numbers. The very small and not exactly measurable C_t values had only the importance of a correction factor used in the calculation of C_f from the experimentally found C values at a few sol concentrations.

In this section we shall on the contrary be interested only in the true reversal of charge concentration C_t. For this can be considered as a measure of the affinity of the cations towards the ionised groups of the colloids. The general equation $C = C_t + C_f$ approximates to the form $C = C_t$ if C_f is very small compared to C_t. This is practically always the case for mono and divalent cations[2] at the small sol

---

[1] P. H. TEUNISSEN and H. G. BUNGENBERG DE JONG, *Kolloid-Beihefte*, **48** (1939) 33.

[2] Except the divalent uranyl cation in the case of phosphate colloids, which cation is fixed very strongly on the ionised groups (see e.g. Fig. 13, p. 283), thus showing relatively low C_t values.

concentrations used in the reversal of charge determinations and even in many cases of trivalent cations the difference between C and C_t is still small.

Thus already the determination of the reversal of charge concentrations at one sol concentration gives us the true reversal of charge concentrations for mono and divalent cations, and the experimental material thus obtained may serve as a basis for the discussion of the relative affinities of these cations for the ionised group of the colloid considered.

For tri and quadri valent cations however, the difference between C and C_t in general already exceeds the experimental errors, so that strictly speaking the exact value of C_t can only be obtained from reversal of charge determination at at least two sol concentrations (see La in Fig. 4 on p. 265).

For the purpose of *comparing* affinities of different cations the absolute values of C_t will not play a rôle in the following discussions, but only their relative magnitudes.

It is easily shown that for this pupose the absolute values of C_t for tri, quadri and even hexa valent cations need not be known, but that the gross reversal of charge concentrations obtained at only one definite sol concentration may serve as well as in the case of mono and diavalent cations, as a basis for discussion.

Writing for three different cations (suffixes 1, 2 and 3) the general equation:

$$C_1 = C_{t_1} + C_{f_1} \qquad C_2 = C_{t_2} + C_{f_2} \qquad C_3 = C_{t_3} + C_{f_3},$$

we may safely assume that at one sol concentration the values of C_{f_1}, C_{f_2} and C_{f_3} are equal (compare the same slope of the hexol and La curve in Fig. 4 on p. 265). From this it follows that if $C_1 < C_2 < C_3$ then also $C_{t_1} < C_{t_2} < C_{t_3}$.

Thus the sequence of the cations 1, 2 and 3 is not altered if we compare at one sol concentration the values of the gross reversal of charge concentrations instead of the true reversal of charge concentrations.

In the various figures which follow we must therefore always remind the reader, that the reversal of charge concentrations of 6, 4 and often 3 valent cations are no longer true reversal of charge concentrations, and that their positions depend on the sol concentration chosen in the experiment. Nevertheless their relative positions are the same as those of the corresponding true reversal of charge concentrations.

b. Complete colloid films at particle surfaces and their use in determining reversal of charge concentrations

In many cases the salts used flocculate or coacervate the sol in a concentration range around the reversal of charge point. In these cases the electrophoretic velocity (at 1/5 depth of the cuvette) is determined as a function of the salt concentration on small flocculi or small coacervate drops.

If however the salt used does not flocculate or coacervate the sol, a method discussed by ABRAMSON[1] for protein sols, proved very useful for our pupose.

Finely divided particles (*e.g.*, SiO_2 particles) are introduced into the sol in which they become covered with a colloid film. By measuring the electrophoretic velocities (at 1/5 depth of the cuvette) of the particles as a function of the salt concentration the reversal of charge concentration can be determined in quite the same

[1] H. A. ABRAMSON, *Elektrokinetic Phenomena*, New York 1934, see p. 147.

way as from measurements on floccules or coacervate drops. In this method the natural phase boundary between the colloid-rich phase (floccules, coacervate drops) and the colloid-poor phase (equilibrium liquid, being a diluted sol) is here replaced by an analogous boundary between an adsorbed "complete" colloid film and the sol in equilibrium with it.

A "complete" colloid film will be said to be present on the surface of a particle, if its electrophoretic properties are solely determined by the sol in which the particle is immersed. Such complete colloid films wholly mask the electrophoretic properties of the original phase boundary of the particle.
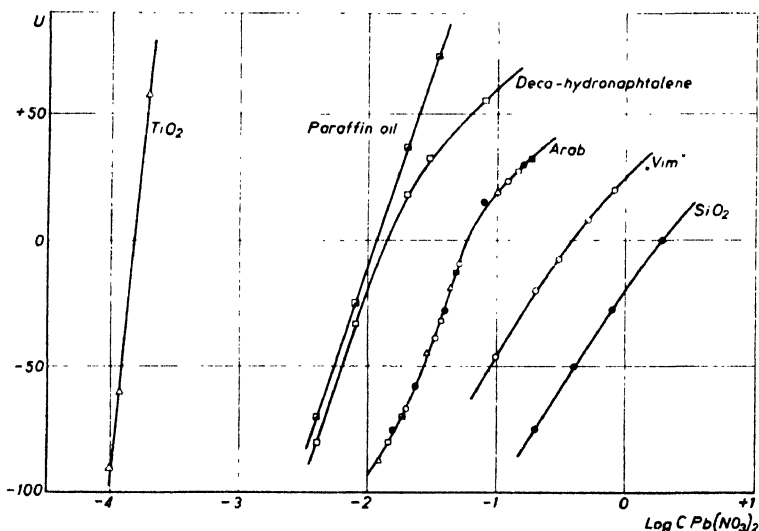
Fig. 6. Reversal of charge of 0.09% Na arabinate sol by Pb(NO₃)₂ determined with the aid of five electrophoretic indicators (see curve "Arab"), together with the reversal of charge of each of these indicators in the absence of Na arabinate.
U = electrophoretic velocity of the suspended particles or drops expressed in arbitrarily chosen units.
log C = logarithm of the Pb(NO₃)₂ concentration in equiv. per l.
Although TiO₂, paraffin oil, decahydronaphtalene, "Vim" (a finely powdered mineral used as an abrasive) and SiO₂ are reversed in charge at very different Pb(NO₃)₂ concentrations, when suspended in 0.09% Na arabinate sol they give the same U — log C curve.

Fig. 6 gives an example, from which may be concluded that with the sol concentration chosen, the five different kinds of "particles" used as "electrophoretic indicators" are covered with a complete colloid film. To avoid misinterpretation we must add that the adjective "complete" does not give any information on the structure of the colloid film formed. Thus complete does not mean, for instance that a monomolecular colloid film on the particle surface would have been completed. It only means that the mono- di- or perhaps poly-molecular colloid film formed possesses such a boundary with the surrounding sol, that its electrophoretic properties have become independent of those of the original particle surface.

In using this method, care must be taken to choose such a sol concentration as to ensure that the particles are really covered with a complete colloid film. For

Fig. 7. Reversal of charge of SiO$_2$ particles with CaCl$_2$ suspensed in exclusively aqueous medium (A) and in egg lecithin sols of increasing concentrations (B — F).
U: electrophoretic velocity of the SiO$_2$ particles expressed in arbitrarily chosen units.
log C: logarithm of the CaCl$_2$ concentrations in equiv. per l.
The lecithin concentrations amount at B to 0.00006% ; at C to 0.0002% ; at D to 0.0006% ; at E to 0.006% and at F to 0.06%.
Increasing the lecithin concentration from E to F does not displace the curve further to the left, i.e., the lecithin concentration at E and F is sufficiently great for the complete covering of the SiO$_2$ particles with a lecithin film.
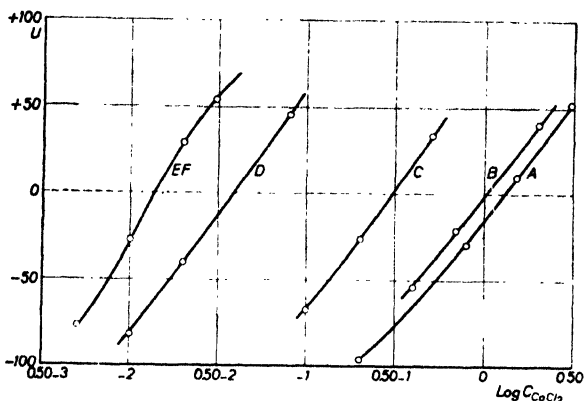
a given kind of particle (*e.g.*, SiO$_2$) the latter may be assumed to be the case when further increase of the sol concentration does not alter the reversal of charge concentration, see Fig. 7 and 8 (left graph). The minimum sol concentrations required
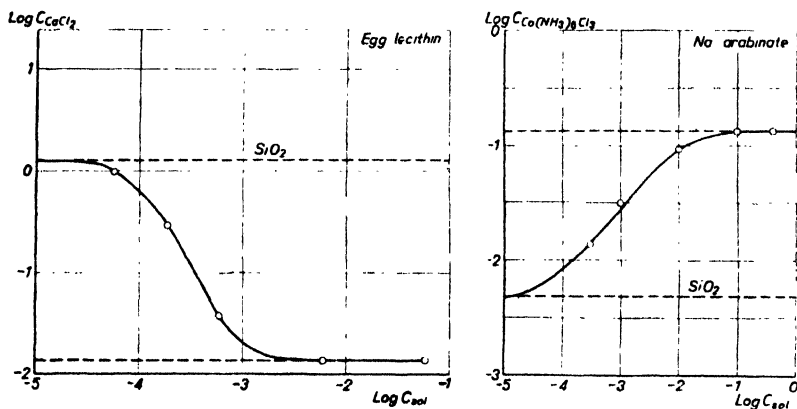
Fig. 8. Coating curves of SiO$_2$ particles with an egg lecithin or with an arabinate film.
Ordinates: logarithm of the CaCl$_2$ or Co(NH$_3$)$_6$Cl$_3$ reversal of charge concentration in equiv. per l.
Abscissae: logarithm of the sol concentration in % (thus 0 = 1% ; —1 = 0.1%, —2 = 0.01% etc.).
Left hand graph.
The CaCl$_2$ reversal of charge concentrations are taken from Fig. 7 (intersections of the U — log C curves with the level U = 0). Since the reversal of charge concentration of the completely naked SiO$_2$ particles lies at a much higher concentration, the coating curve here falls from the upper dotted level to the lower dotted level on increase of the sol concentration.
Right hand graph.
According to data from a graph similar to that of Fig. 7 but in which the displacement of the U—log C curves on increase of the sol concentration is just the opposite. This is a consequence of the reversal of charge concentration of completely naked SiO$_2$ particles lying at a much lower concentration of the Co(NH$_3$)$_6$Cl$_3$ than that of the particles coated with arabinate. Consequently the coating curve here rises on increase of the sol concentration from the lower dotted level to the upper dotted level.

differ from colloid to colloid but are in general relatively low, ranging usually between 0.1 and 0.001%; compare the two graphs of Fig. 8. For particulars see the original publication [1].

Faced with the plan of determining the reversal of charge concentrations with a large number of cations, it is an immense task to investigate for each cation the minimum sol concentration needed. For lack of time it was limited to only one cation. But to ensure a margin of safety, in determining the reversal of charge concentration with the remaining cations, the sol concentration was chosen much higher (*e.g.*, ten-fold) than the minimum concentration so obtained.

The formation of colloid films is not restricted to the suspended SiO_2 particles, but occurs also on the glass walls of the electrophoretic cuvette. If the film on the glass walls is complete no electroendosmosis will occur along this film at the reversal of charge point of the colloid. A peculiar consequence is, that then measurements at 1/5 depth (giving true velocities) and on 1/2 depth (giving velocities distorted by electroendosmosis) lead nevertheless to the same reversal of charge concentration. Fig. 9 gives an example.
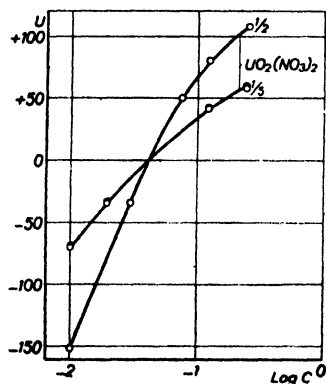
Fig. 9. Electrophoretic velocity of SiO_2 particles suspended in a 0.1% Na arabinate sol as a function of the uranyl nitrate concentration, measured at $1/5$ and at $1/2$ cell height.
U: electrophoretic velocity of the SiO_2 particles expressed in arbitrarily chosen units.
log C: logarithm of the $UO_2(NO_3)_2$ concentration in equiv. per l.
$1/5$ and $1/2$: measurements at $1/5$ and at $1/2$ cell height. Not only the SiO_2 particles but also the glass wall of the electrophoresis cell is completely covered by an arabinate film. As a result of this the electroendosmosis along the glass wall is suppressed at the reversal of charge point of the arabinate and the two U—log C curves thus intersect the level O at the same $UO_2(NO_3)_2$ concentration.

c. Phosphate colloids

Figs. 10 and 11 give for two different preparations of egg lecithin the interpolation graphs used for determining the reversal of charge concentration of a number of mono and divalent cations. (Li, Na, K, Mg, Ca, Sr and Ba as chlorides, the rest as nitrates).

It appears, that a considerable spread of the curves occurs, every cation behaving specifically.

[1] P. H. TEUNISSEN and H. G. BUNGENBERG DE JONG, *Kolloid-Beihefte*, 48 (1938) 33. Among the colloids studied the lowest values (about 0,001%) were encountered among phosphatides and nucleates.
L. TEUNISSEN-VAN ZIJP, *Thesis*, Leiden 1938, studied gelatin and clupein and found complete covering of SiO_2 particles to occur at 0.001—0.003%.

Fig. 10. Reversal of charge of egg lecithin (purified via the CdCl$_2$ compound) with alkali and alkaline earth chlorides.

Since the purified egg lecithin does not flocculate with the salts used, the electrophoresis measurements were carried out on suspended SiO$_2$ particles, which at the chosen sol concentration (0.049%) were covered with a complete colloidal film (this is even the case from a 10 × smaller lecithin concentration onwards (see Fig. 8).

U: electrophoretic velocity of the SiO$_2$ particles expressed in arbitrarily chosen units.

log C: logarithm of the salt concentration in equiv. per l. (thus 0 = 1 N; —1 = 0.1 N; —2 = 0.01 N; —3 = 0.001 N).

The intersections of the curves with the dotted level (U = 0) give the reversal of charge concentrations. It is only with KCl that this cannot be reached since in that case it lies higher than that of the saturated solution.

The sequence of the cations is with the exception of one detail (interchange of Ba and Sr) the same as in Fig. 11. Only the reversal of charge concentrations lie lower, to which we return again on page 294.
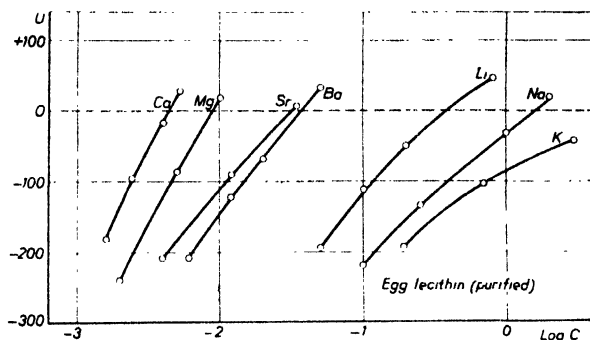
Fig. 11. Reversal of charge of egg lecithin sensitised with cholesterol with a number of chlorides (Li, Na, K, Mg, Ca, Sr, Ba) or nitrates (the remaining cations).

The egg lecithin (commercial preparation) was first reasonably freed from impurities by repeated precipitation with acetone from ether solutions. Since a clear transparent sol prepared from this does not flocculate with salts or does so extremely slowly, and we wished to determine the electrophoretic velocities on the floccules themselves, a sol containing 50% cholesterol (50% with respect to lecithin) was prepared.

The cholesterol (like for that matter numerous other organic non-electrolytes such as triolein) sensitises the sol, i. e., so changes it that it flocculates rapidly with electrolytes.

U: electrophoretic velocity of the floccules expressed in arbitrarily chosen units.

log C: logarithms of the salt concentrations in equiv. per l. (thus 0 = 1 N; —1 = 0.1 N; — 2 = 0.01 N etc.).

The intersections with the level U = 0 give the reversal of charge concentrations (for KCl unreachable since this lies higher than that of the saturated solution).
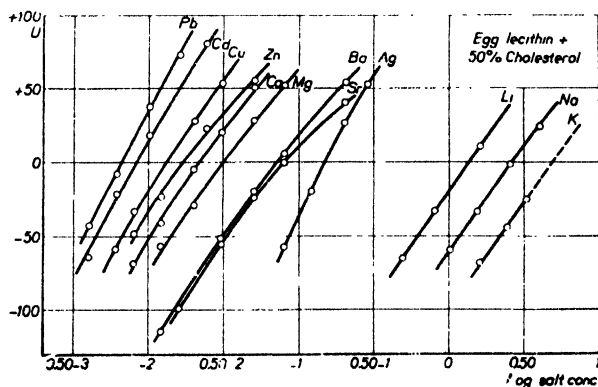
Regularities in the cation sequences may be best surveyed by diagrams in which the reversal of charge concentrations each time for a group of related ions are put together. Such a diagram — which we will call a "reversal of charge spectrum"

or a "cation spectrum" of the colloid considered — is given in Fig. 12, which not only contains the reversal of charge concentrations for the mono and divalent cations of Fig. 11, but also for UO_2, Mn, Co, Ni and higher valent cations.

In Fig. 13 similar cation spectra of three other phosphate colloids (two different fractions of soya bean phosphatides and sodium nucleate) are given.
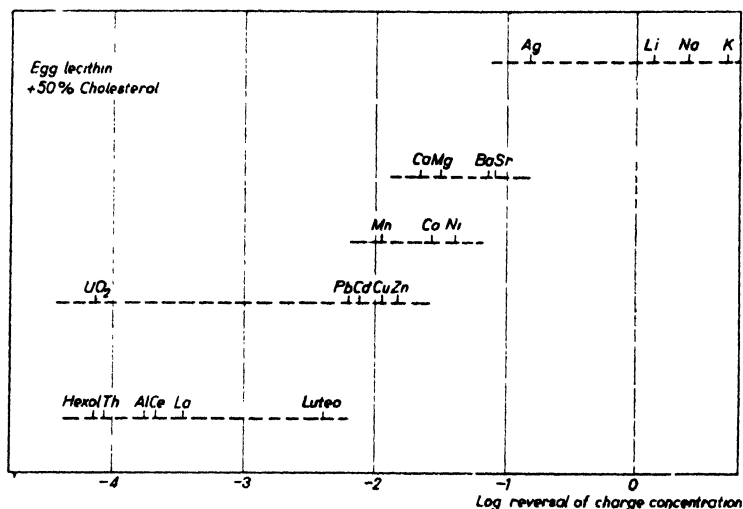
Fig. 12. Reversal of charge spectrum of egg lecithin (sensitised with cholesterol).
For the sake of surveyability the logarithms of the reversal of charge concentrations (can be read off from Fig. 11 and supplemented with similar data for UO_3, Mn, Co, Ni and cations of higher valency) are plotted not on one level but on several horizontal levels. The uppermost level contains data on monovalent cations, the succeeding three data on divalent cations and the lowest data on cations of higher valency.

The following points may be considered as characteristic of phosphate colloids:

1. Increasing the valency of a cation of the A subgroups of the Periodic System decreases considerably the reversal of charge concentration. Compare the relative positions of the monovalent Li, Na, K with the divalent Mg, Ca, Sr, Ba and with the trivalent Ce, La.

2. Increasing the ion radius within the above named mono and trivalent ions (thus Li → Na → K or Ce → La) increases the reversal of charge concentration. This does not apply to the divalent ions (Mg, Ca, Sr, Ba) in which irregular series appear. The relatively small Ca ion has here often the lowest reversal of charge concentration. These irregular series will be discussed further in § 2 g (p. 288).

3. Ions of the B subgroups of the Periodic System show in general smaller reversal of charge concentrations than those of the A subgroups with equal valency. Compare Ag, Tl with Li, Na,K and Pb, Cd, Cu, Zn with Mg, Ca, Sr, Ba.

4. UO_2, though divalent, shows an extraordinarily low reversal of charge concentration.

Before discussing these results, we shall first consider the reversal of charge behaviour of carboxyl and sulphate colloids.
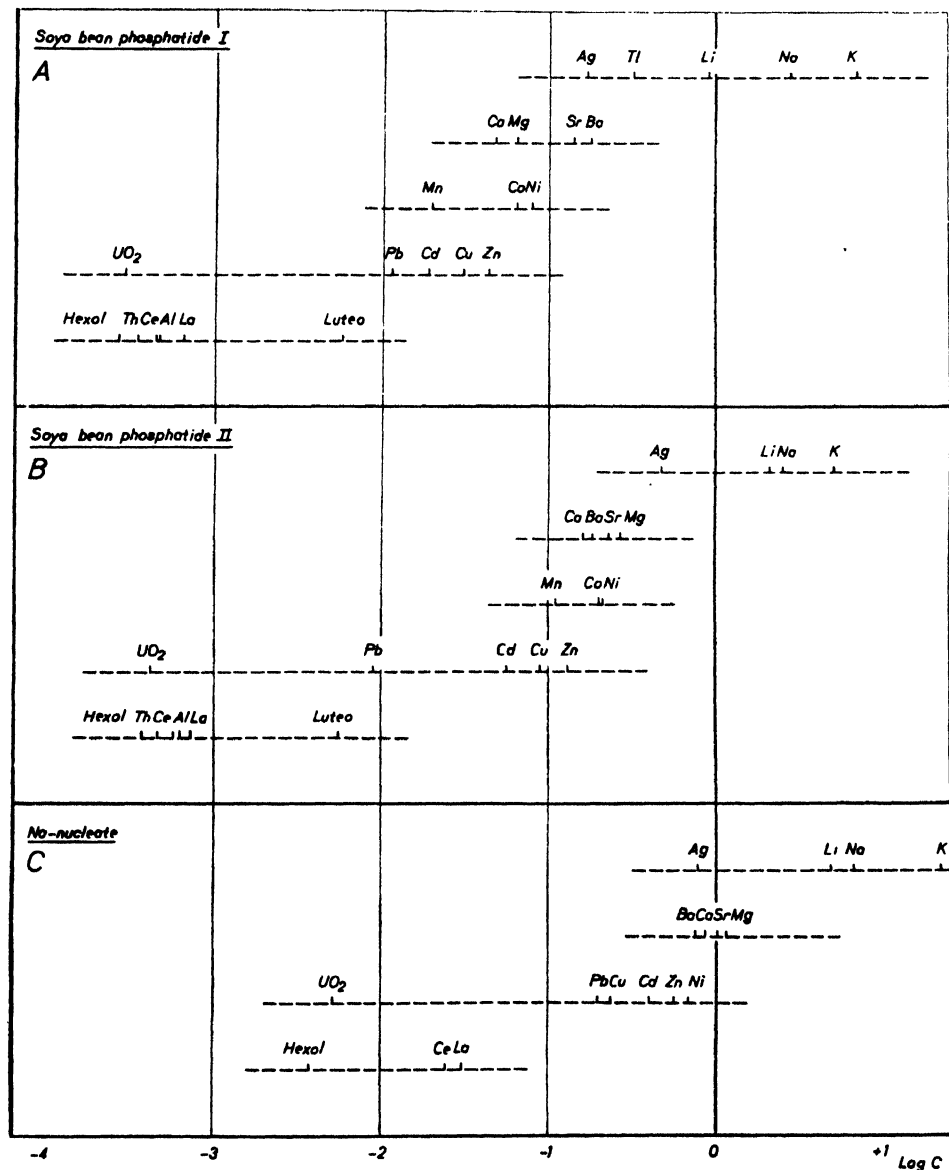
Fig. 13. Reversal of charge spectra of some other phosphate colloids.

A: of alcohol soluble soya bean phosphatide, sensitised with triolein.

B: of alcohol insoluble soya bean phosphatide measured directly on the floccules, except for Li, Na and K where the measurements were carried out on suspended SiO_2 particles.

C: of Na nucleate, measurements on floccules, except for Li, Na and K, where the measurements were made on suspended SiO_2 particles.

d. Carboxyl colloids

Fig. 14 gives the ion spectra for three carboxyl colloids. The data for the third one, Na-pectate, are very incomplete because of technical difficulties.
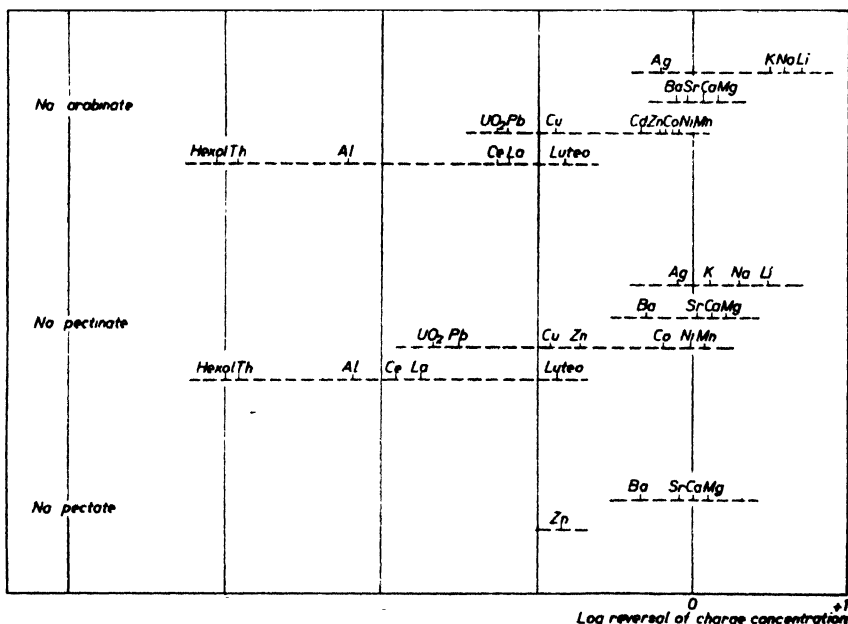
Fig. 14.   Reversal of charge spectra of some carboxyl colloids (see text).
On account of great technical difficulties only a very restricted number of cations were investigated with Na pectate.

The following points are characteristic:

1. Increasing the valency of a cation of the A subgroups of the Periodic System decreases the reversal of charge concentration, though this influence is less pronounced than in the case of phosphate colloids. Compare: Li, Na, K with Mg, Ca, Sr, Ba and with Ce, La.

2. Increasing the ion radius within the above named groups has a different effect. Only in the case of the trivalent ions is the sequence still the same as in phosphate colloids (Ce⟨La). In the mono and divalent cations increasing the ion radius decreases the reversal of charge concentration (sequences K⟨Na⟨Li and Ba⟨Sr⟨Ca⟨Mg in Fig. 14).

3. Ions of the B subgroups of the Periodic Systems show in general smaller reversal of charge concentration than those of the A subgroups with the same valency (compare Ag with Li, Na, K and Pb, Cd, Cu, Zn with Mg, Ca, Sr, Ba).

4. $UO_2^{..}$ occupies no exceptional position (as in phosphate colloids) in showing a reversal of charge concentration not much smaller than Pb$^{..}$.

Comparing the results of carboxyl colloids with those of phosphate colloids, we may state that in the points 1 and 3 the resemblance is close, and that only in the points 2 and 4 is a different behaviour met with.

e. Sulphate colloids

Fig. 15 gives the cation spectra of three sulphate colloids. In the upper spectrum the usual monovalent ions are not recorded, as they have not been determined. They have been determined in media containing 40% acetone, and on the floccules then obtained the same order (Ag $<$ K $<$ Na $<$ Li) was obtained as shown in the two other ion spectra.
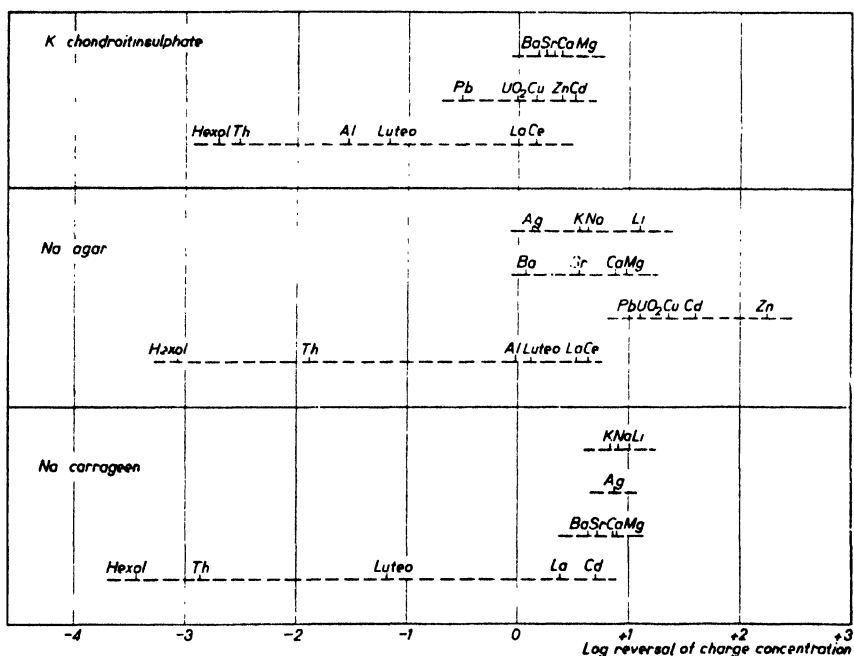
Fig. 15. Reversal of charge spectra of some sulphate colloids (see text).

The following points are characteristic for sulphate colloids:

1. Increasing the valency of a cation of the A subgroups of the Periodic System does not decrease at all strikingly the reversal of charge concentration. Compare Fig. 15, in which the groups Li, Na, K and Mg, Ca, Sr, Ba and Ce, La lie nearly in the same range of concentrations.
2. Increasing the ion radius within the above named groups has decreasing effect on the reversal of charge concentration (K $<$ Na $<$ Li and Ba $<$ Sr $<$ Ca $<$ Mg and La $<$ Ce).
3. Ions of the B subgroups of the Periodic System no longer systematically show a smaller reversal of charge concentration than those of the A subgroups with

the same valency. In many cases they fall in the same order of magnitude (for instance Ag in the case of carrageen; UO_2, Cu, Cd, Zn in the case of chondroitin sulphate) or even at much higher concentration than those of the A subgroups (Pb, UO_2 ,Cu, Cd, Zn in the case of Agar).

4. As in carboxyl colloids $UO_2^{..}$ shows no exceptional position with respect to the divalent ions of the B subgroups of the Periodic System.

Comparing the results with those of the phosphate colloids, we may state, that in all four points the reverse is obtained. The differences with the carboxyl colloids are also considerable, as only point 4 and the sequence of the alkali — and alkaline earth cations in point 2 are alike. Still we may consider the ion spectra of the carboxyl colloids as intermediate between the two extremes of phosphate and sulphate colloids, though the carboxyl colloids stand much nearer to the phosphate colloids than to the sulphate colloids.

One of the most startling results with the sulphate colloids is certainly the first point of the above survey, which formulates the incapacity of the cation valency to decrease markedly the reversal of charge concentration. This is the more startling as from the study of the electroviscous effect of the agar sol it seemed evident that the cation valency was the factor of primary importance in discharging (see p. 203 Ch. VII § 5b and Fig. 16). Therefore electrophoretic measurements have been made on agar in a wide concentration range, comprising the range of small concentrations (0—10 m. eq. p. l) in which the suppression of the electroviscous effect is accomplished for the greater part. Fig. 16 shows the results.
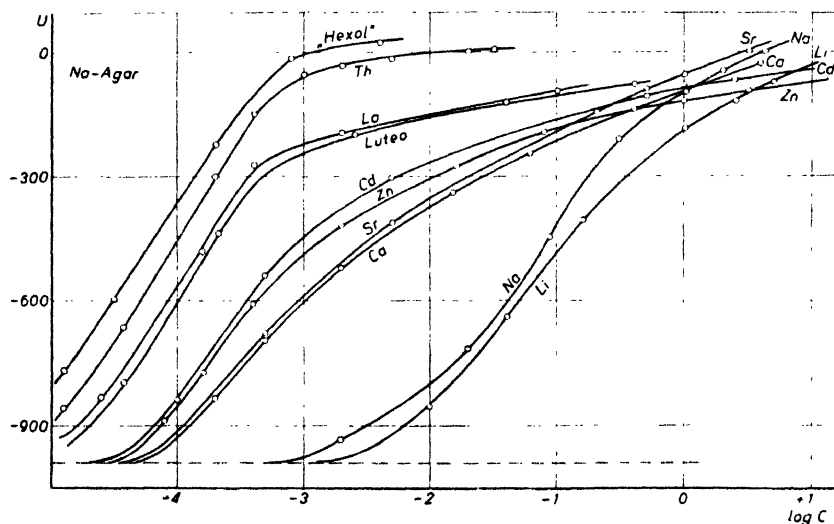
Fig. 16. Electrophoretic velocity of Na agar in dependence on the concentration of added salts (measurements on suspended SiO_2 particles).
U: electrophoretic velocity expressed in arbitrarely chosen units.
In the region of smaller salt concentrations (up to log C = — 1, i.e., 0.1 N) it is the valency of the cation principally which determines the sequence of the curves. This regularity disappears at higher concentrations (see text).

From the figure it is seen that indeed in the range of small concentrations (log C = — 4 = 0,1 m. eq. p. l; — 3 = 1 m. eq. p. l; — 2 = 10 m. eq. p. l) the cation valency is the factor of primary importance in lowering the electrophoretic velocity. At higher concentrations however this regularity is lost and the curves intersect one another in various ways before the reversal of charge concentrations are reached. They lie in many cases so high that they can be only estimated by extrapolation (compare for instance in Fig. 15 values of + 2, corresponding to the quite fictitious concentration of 100 N!).

From this different behaviour of agar towards valency influences in small and in large concentrations, we may already gain indications, that the mechanism which leads to reversal of charge is of a quite different nature than that which occurs at low concentrations.

This latter is relatively unspecific, and consists only in a replacing the original counter ions by the cations of the added salt and in a further "compressing of the double layer". At higher concentrations preceding the reversal of charge, the cations must be fixed on the ionised groups of the colloid and now specific properties both of ions and of the ionised groups are of primary importance. This may explain why in the suppression of the electroviscous effect all three kinds of colloids behave nearly in the same manner, and why in the reversal of charge phenomenon very marked specific differences are met with.

f. The rôle of the polarisability of the ionised groups.
Comparison of the cation spectra of phosphate and sulphate colloids

The characteristic differences which exist between the cation spectra of the above three groups of colloids can be explained by assuming the following order of polarisability for the ionised groups and for water:

phosphate group $>$ carboxyl $>$ H_2O $>$ sulphate group.

This assumption seems in accordance with approximate calculations for which we refer to the original publication[1].

The affinity between oppositely charged ions depends not solely on the valency of these ions but several factors play a rôle.

If our assumption is true that the reversal of charge is generally caused by fixation of a sufficient amount of cations on the ionised groups analogous to the experimentally well-founded case of hexol ions (see p. 262 § 1 b) then the affinity of cations and ionised groups must depend on *valency, radius and polarising power of the cation and on the polarisability of the negatively charged ionised group of the colloid*. But as the phenomenon of fixation of cations takes place in aqueous medium the *polarisibility of water* itself must also play a rôle. As to both polarisabilities, we may say that it is most important which of the two is the greater one; that of the ionised groups or that of water. We shall show that the on almost all points characteristically opposite behaviour of phosphate colloids and sulphate colloids can in fact be explained easily from the different polarisabilities of the ionised groups, the phosphate group being stronger, the sulphate group being less polarisable than water.

If we consider for instance the monovalent ions Li, Na and K, then — exclusively from the point of view of "field strength" — on the surface of these ions the fixation on a given negatively charged ionised group will be easiest in the case of the smallest

---

[1] P. H. TEUNISSEN and H. G. BUNGENBERG DE JONG, *Kolloid-Beihefte*, 48 (1939) 33, see p. 80.

ion— Li—, and will be increasingly more difficult for the larger Na and K ions. But as a second influence we must consider the energy of polarisation. If the ionised group is more polarisible than water, then the polarisation energy is added to the Coulomb energy. In this case the above order of cations will not be disturbed, on the contrary the spread of Li — Na — K will possibly be strengthened.

For the smallest ion in question (Li) after it has lost part of its hydration water will add on fixation more polarisation energy than the greater Na and K ions. Thus for an ionised group more polarisable than water the order of decreasing affinity will be Li $>$ Na $>$ K. And as the reversal of charge concentrations can be considered as an inverse measure of cation affinity we must expect that the reversal of charge concentrations increase in the order Li $<$ Na $<$ K.

This indeed is the characteristic order found in phosphate colloids (p. 281-283, Fig. 10-13).

Considering now the same cations in the case that the polarisability of the ionised group is less than that of water, then the polarisation energy of the water molecules (cation hydration) will oppose the Coulomb energy. Now the Li ion, with smallest radius and therefore the most hydrated of the three will be fixed most difficultly on the ionised group. It will have the greatest tendency to rest in the perfect hydrated state free in the medium.

Therefore the affinity order will be K $>$ Na $>$ Li, and the order of the reversal of charge concentrations K $<$ Na $<$ Li, which is indeed found in the case of sulphate colloids (p. 285, Fig. 15).

As in the physical chemistry of electrolytes the solubility of salts possessing different cations but the same anion depends on the same factors as used in the above discussion *viz.*, valency, radius and polarising power of the cation and polarisability of the anion, it may be expected, that cation sequences may occur similar to the cation sequences found in the reversal of charge concentrations of colloids. This indeed is the case, so the solubilities of phosphates (moles in 1 000 moles H_2O) increases in the order Li $<$ Na $<$ K, however of sulphates in the order K $<$ Na $<$ Li.

In discussing the remaining characteristic differences in the cation spectra of the three classes of colloids, we shall however not repeatedly discuss such analogies in the solubility of ordinary salts. Therefore the original publication may be consulted.

It is however interesting to note that in one case (the position of Ag in certain sulphate colloids) in which theoretically a difficulty in explanation exists, this difficulty also exists in the relatively small solubility of silver sulphate. See § 2 h, p. 290.

g. Comparison of the cation spectra of phosphate and sulphate colloids continued. Influence of the ion radius. Abnormal sequences

If we continue the discussion of the influence of the radius of in other respects comparable cations (having equal valency and belonging to the A subgroups of the Periodic System) we might expect on the same reasoning as in § 2 f, for the divalent ions the following sequences for the reversal of charge concentrations:

$\quad$ Mg $<$ Ca $<$ Sr $<$ Ba $\quad$ for phosphate colloids $\quad$ **A)**

$\quad$ Ba $<$ Sr $<$ Ca $<$ Mg $\quad$ for sulphate colloids $\quad$ **B)**
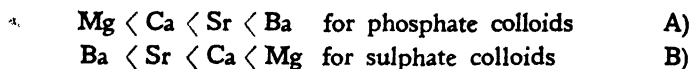
and for the trivalent ions[1][2] :

---

[1] The position of the trivalent Al ion, is discussed later, see p. 291, § 2 i.

[2] The radius of the Ce ion is smaller than that of La and is caused by the so-called "contraction of the Lanthanides".

$$Ce < La \qquad \text{for phosphate colloids} \qquad C)$$

$$La < Ce \qquad \text{for sulphate colloids} \qquad D)$$

Of these four expectations the last three are in accordance with the experimental findings, the first one however is not fulfilled. In phosphate colloids (p. 281-283, Fig. 10-13 and p. 295, Fig. 20) we met with various sequences:

$$Ca < Mg \lll Sr < Ba$$
$$Ca < Mg \lll Ba < Sr$$
$$Ca < Ba \;\; < Sr < Mg$$
$$Ba < Ca \;\; < Sr < Mg$$

standing between the two extreme sequences:

$$Mg < Ca \;\; < Sr < Ba$$
$$\text{and} \;\; Ba < Sr \;\; < Ca < Mg$$

and which therefore may be called *transition sequences.*

Such transition series have later[1] been shown also to occur in phosphate colloids with the alkali cations, if we not only investigated Li, Na and K but continued this series with the still larger ions Rb and Cs. The sequences found are the following: (see also Fig. 17)
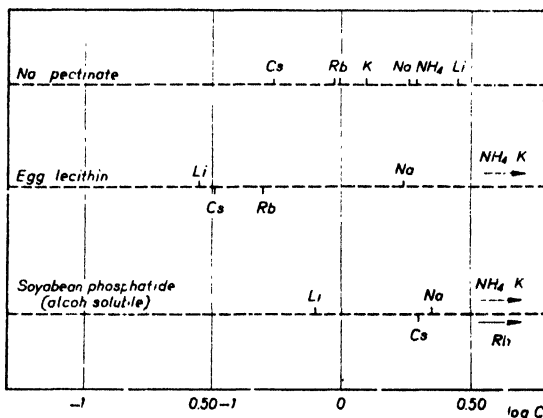
Fig. 17. Reversal of charge spectra of Na pectate (carboxyl colloid) and of egg lecithin and alcohol soluble soya bean phosphatide (phosphate colloids) with alkali chlorides (and NH_4Cl). The alkali cations form transition series with the two phosphatides (see text).

phosphate colloids $\left\{ \begin{array}{l} Li < Cs \; < Rb < Na < K \quad \text{egg lecithin} \\ Li < Cs \; < Na < Rb < K \quad \text{soya bean phosphatide (alcohol-} \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{soluble fraction)} \end{array} \right.$

carboxyl colloid $\qquad Cs < Rb < K \;\; < Na < Li \quad$ Na-pectate

sulphate colloid $\qquad Cs < Rb < K \;\; < Na < Li \quad$ Na agar

from which it may be seen that in sulphate and carboxyl colloids the position of Cs and Rb is quite as might be expected (see p. 284 and 288), but that in phosphate colloids seemingly abnormal sequences are obtained.

In this case of the alkali cations they are amenable to simple explanation, from which also the denomination "transition series" appears to be adequate.

---

[1] H. G. BUNGENBERG DE JONG, L. TEUNISSEN-VAN ZIJP and P. H. TEUNISSEN, *Kolloid-Z.*, 91 (1940) 311.

Writing both abnormal sequences in the following way:

$$Li \longrightarrow Na \longrightarrow K \qquad\qquad Li \longrightarrow Na \longrightarrow K$$

$$\text{and}$$

$$Cs \longleftarrow Rb \qquad\qquad Cs \longleftarrow Rb$$
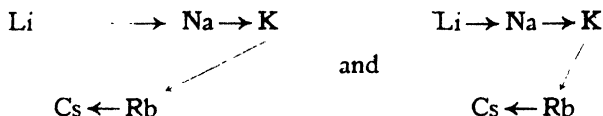
the influence of the polarisation energy shows itself quite as may be expected. For as the phosphate group is more polarisable than water, the smallest ions will show, as we already explained in § 2 f, the sequence $Li < Na < K$.

But the very large Cs and Rb ions have, because of their large radius, too small a field strength to show polarising action. For these ions the sequence $Cs < Rb < K$ must thus be expected.

The experimentally found irregular sequences can thus be explained as the result of the still further increase of the cation radius.

It seems therefore logical that the irregular sequences found in phosphate colloids with the divalent cations might be explained on similar points of view. Still unknown factors must also play a rôle, for it is not possible to write these sequences in the way we did above for the alkali cations. See also p. 294.

h. Comparison of the cation spectra of phosphate and sulphate colloids continued. Cations of the B subgroups compared with cations of the A subgroups

Cations of the B subgroups of the Periodic System have, as is well known, a stronger polarising power than cations of the same valency belonging to the A subgroups. Therefore it might be expected that in phosphate colloids their reversal of charge concentrations are smaller, in sulphate colloids larger than those of the A subgroups.

Indeed in phosphate colloids (Fig. 12 and 13, p. 282 and 283) we meet the sequences:

$$Ag, Tl < \text{alkali cations}$$
$$UO_2, Pb, Cu, Cd, Zn < \text{alkaline earth cations.}$$
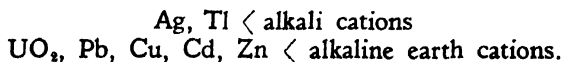
In sulphate colloids the expected reversal position tends to be present or is actually present in the case of the divalent ions, the reversal of charge concentrations of the ions of the B subgroups lying at the same order of concentration as those of the A subgroups (chrondroitin sulphate, carrageen) or actually at higher concentrations (agar). See p. 285, Fig. 15.

Contrary to the expectation Ag shows in chondroitin sulphate and agar a still lower reversal of charge concentration than the alkali cations. It is remarkable, that the analogy with the solubility or salts is very close, for the solubility of Ag_2SO_4 is also smaller than that of K_2SO_4.

In carrageen the order $K < Ag < Na < Li$ is found, which comes thus somewhat nearer to the expected sequence $K < Na < Li < Ag$.

The very low reversal of charge concentration of $UO_2^{..}$ characteristic of phosphate colloids and not found in carboxyl or sulphate colloids, finds its analogy in the relative solubilities of uranyl phosphate, uranyl acetate and uranyl sulphate.

i. Comparison of the cation spectra of phosphate and sulphate colloids continued. Influence of the cation valency. The individual behaviour of complex ions, of Th and Al

Now turning to the valency of the cation, it must be expected that in phosphate colloids increase of valency would decrease the reversal of charge concentration, in sulphate colloids however the contrary should occur.

Restricting ourselves to mono, di and trivalent cations of the A subgroups (Li, Na, K; Mg, Ca, Sr, Ba; La, Ce), the expectation is once more fulfilled in phosphate colloids, in which a considerable spreading of reversal of charge concentrations in the expected order is met with.

In sulphate colloids, the tendency to the reversal sequence shows itself in the fact, that the groups of mono, di and trivalent cations of the A subgroups lie in nearly the same range of concentrations. In every case, the very marked valency influence shown in phosphate colloids has disappeared here altogether.

The marked difference in ion spectra of phosphate and sulphate colloids as regards the relative positions of the reversal of charge concentrations for mono, di and trivalent cations, is no longer found in the tetravalent Th ion, this cation showing low reversal of charge concentrations in sulphate colloids as well as in phosphate colloids. More or less the Al cation too does not behave like the other two trivalent cations La and Ce, it having also the tendency to cause reversal of charge at relatively low concentrations. This individual behaviour of Th and of Al can be explained by assuming that these cations are not present in water as such, but as definite complex ions (the metal ions being surrounded by a definite number of water molecules) or still larger complexes originating from hydrolysis. As such large complex ions have of course no longer any polarising power and thus the effects of it — e.g., in sulphate colloids, loss of the valency influence on the reversal of charge concentration — are also absent.

Thus Th and Al may be more or less compared with the two complex ions hexol and $Co(NH_3)_6^{...}$, which also always show a marked relative valency influence, the former hexavalent one always causing reversal of charge at very low concentrations, the latter trivalent one at higher concentrations.

j. The cation spectra of carboxyl colloids

After having discussed at some length the cation spectra of phosphate and sulphate colloids, we can be brief in discussing those of carboxyl colloids.

The following points (compare p. 284, Fig. 14; p. 289, Fig. 17) are in accordance with our assumption that the ionised carboxyl group is more polarisable than water:

a. the influence still present of the cation valency (A subgroups) in lowering the reversal of charge concentration, though not so marked as in phosphate colloids.

b. the lower reversal of charge concentrations of cations belonging to the B subgroups of the Periodic System, compared to those of the same valency belonging to the A subgroups.

c. The sequence Ce < La

There are however also points, showing that the polarisability of the ionised carboxyl group of these colloids can only be slightly stronger than of water:

a. the cation sequence Cs $\langle$ Rb $\langle$ K $\langle$ Na $\langle$ Li
b. the cation sequence Ba $\langle$ Sr $\langle$ Ca $\langle$ Mg
which are the same as those found in sulphate colloids.

The polarising effects of the relatively voluminous alkali and alkaline earth cations are here evidently no longer sufficient to cause a reversal of the sequence, or a transition series respectively (as in phosphate colloids).

k. Constitutional influences on the polarisability of the ionised group

The results obtained show that from a physico-chemical point of view the division of the biocolloids considered according to the composition of the ionised groups into phosphate, carboxyl and sulphate colloids is a fertile one. Their behaviour depends not only on their equivalent weight, but also on the said composition.

Nevertheless the discrimination into the said three groups of colloids is only a beginning, for there are many indications that the polarisability of the ionised groups may be altered by constitutional influences.

This is clearly seen by comparison of the ion spectra of Fig. 14 (p. 284), with that for ammonium oleate, see Fig. 19, in which only as many cations were investigated (Fig. 18) as would suffice to characterise it.
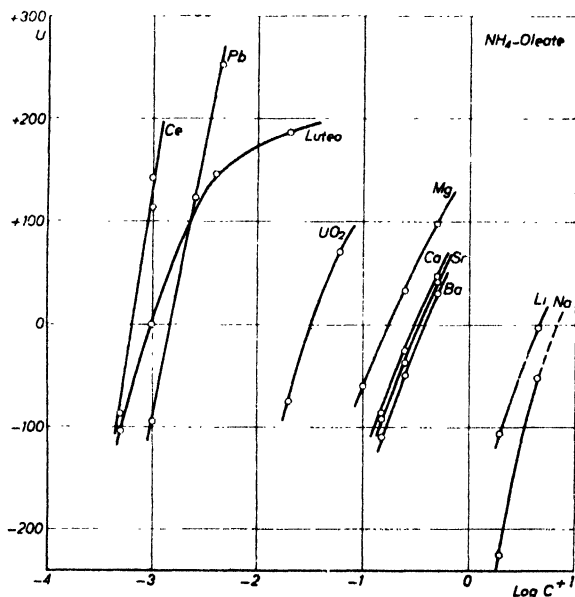
Fig. 18. Reversal of charge of ammonium oleate.
U = electrophoretic velocity expressed in arbitrarely chosen units. The reversal of charge point cannot be reached with NaCl.
This also holds for KCl, for which the measurable portion of the U — log C curve lies so far below the NaCl curve that this part of the curve falls outside the figure.

In this oleate spectrum (Fig. 19) all points characteristic of a ionised group more strongly polarisable than water are met with:

1. The strong influence of ion valence in lowering the reversal of charge concentrations.
2. The ions Pb (B subgroup) and UO_2 have lower reversal of charge concentrations, than those of the alkaline earth.
3. In the alkali and alkaline earth cations the sequence is Li $\langle$ Na $\langle$ K and Mg $\langle$ Ca $\langle$ Sr $\langle$ Ba.
4. UO_2·· does not take a very extreme position (as in the case of phosphate colloids).

Compared with the carboxyl colloids hitherto mentioned (arabinate, pectinate, pectate), the points 1, 2 and 4 are the same.

However the sequences mentioned in 3) are just the reverse. This points to a greater polarisability of the carboxyl group in oleates than in arabinates, pectinates and pectates.

Fig. 19. Reversal of charge spectrum of ammonium oleate. The figure is drawn from the data of Fig. 18. The reversal of charge concentration of K can not be given in this case. K →→ signifies the direction in which the latter lies. The sequence of the alkali and alkaline earth ions is characteristic, precisely the reverse of that with the carboxyl colloids discussed so far. These reversed sequences were also found with Na oleate.
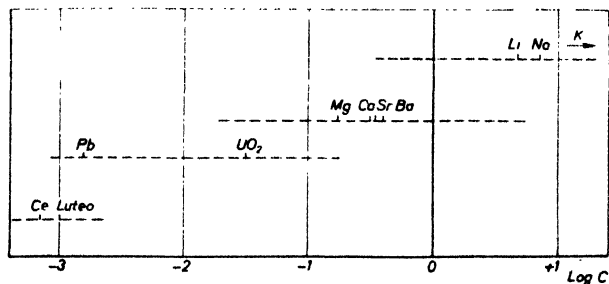
It seems quite natural to ascribe the lesser polarisabilities of the carboxyl group of the latter to a certain constitutional influence. They are derived from polymer carbohydrates and thus the presence of hydroxyl groups in the neighbourhood of the carboxyl group could be the cause of the decreased polarisability (cf. the well known action of hydroxyl substitution on the dissociation contants of fatty acids).

In this connection it is interesting that analogies are found also in salt solubilities. Thus the sequence for benzoates $Li < Na < K$ is altered by introduction of OH to that of salicylate $K < Na < Li$.

The sequence of formates is $Li < Na < K$, and of tartrates is $Na < Li < K$, the latter being a "transition series".

In order to check this supposition, Teunissen, Rosenthal and Zaayer[1] compared the reversal of charge concentrations of oleate, trioxystearate, hexaoxystearate and arabinate for the alkali and alkaline earth cations. In this investigation also another factor — the influence of pH — showed it self very markedly. In the following survey the reversal of charge sequences at different pH values follow one another in such a way, that the inverse sequence for oleate (at pH 10 but also at pH 7) and arabinate (at pH 10 but also at pH 6) are interlinked as successively as possible. The intermediate sequences have in part the character of transition series (§ 2 g);

Oleate pH 7—10	Mg — Ca — Sr — Ba	Li — Na — K
Trioxystearate pH 10	Mg — Ca,Sr — Ba	Li,Na — K
Hexaoxystearate pH 10	Mg — Ca,Ba — Sr	Li,Na — K
Trioxystearate pH 6	Ca — Mg — Sr — Ba	K — Na,Li.
Hexaoxystearate pH 6	Ba — Ca — Sr — Mg	K — Na,Li.
Arabinate pH 6—10	Ba — Sr — Ca — Mg	K — Na — Li.

[1] P. H. Teunissen, S. Rosenthal and W. H. Zaayer, Rec. Trav. Chim. Pays Bas., 57 (1938) 929.

The survey shows that by introduction of OH groups into oleate, the cation sequences change in the direction of those characteristic of arabinate; hexaoxystearate showing at pH 6 already almost wholly the sequences of arabinate.

It may be expected, that also introduction of other so-called "negative groups" (which in general increase also the dissociation constant of weak acids) have a similar influence.

From this point of view probably the transition sequences found with phosphate colloids in the alkaline earth cations (p. 289) may be explained. Compare the Figures 10-13 on p. 281-283.

All facts point to the phosphate group being more polarisable than water. We should thus expect the sequence: $Mg < Ca < Sr < Ba$. Nearest to this comes the sequence found in egg lecithin and soya bean phosphatide (soluble in alcohol) $Ca < Mg \ll Sr < Ba$.

If the occurrence of transition sequences indicates a decrease in polarisability, then this is perhaps caused by the presence of ester groups in the phosphatide molecule, in the immediare neighbourhood of the phosphate group.

In Na-nucleate we meet with the sequence·

$$Ba < Ca < Sr < Mg,$$
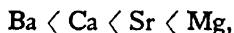
a transition sequence which stands much nearer to the completely inverse sequence $Ba < Sr < Ca < Mg$. But in the nucleate molecule the phosphate group is directly esterified with two sugar residues, from which a considerable decrease of the polarisability may be expected.

For the polarisability of the carboxyl group in proteins see § 3a (p. 297).

1. The absolute values of the reversal of charge concentrations

In the preceeding discussions we used the relative positions of the reversal of charge concentrations of different cations for one and the same colloid as an inverse measure of the relative affinities of those cations for the colloid in question. It is certainly not permissible to consider the absolute values of these reversal of charge concentrations as a direct measure of the cation affinities, for there are indications that these absolute values still depend on further factors.

The available material shows that in phosphatides these absolute values depend on the "density of charge" of the colloid considered[1].

In Fig. 20 a number of ion spectra for different phosphatide preparations are given. They have been arranged in the order of decreasing reciprocal hexol number, the uppermost showing the highest (76 000), the lowest one showing the lowest reciprocal hexol number (782).

Fig. 20 shows, that the characteristic sequence of cations is in all the same (or practically the same), only the spread is greatest with low "charge density", this becoming less with increasing "charge density" (i.e., with decreasing reciprocal hexol number).

We may add that the reversal of charge with H ions (not marked in the figure) also shows the same behaviour, the *I.E.P.* shifting from higher to lower pH values with decreasing reciprocal hexol numbers.

---

[1] H. G. BUNGENBERG DE JONG, L. TEUNISSEN-VAN ZIJP and P. H. TEUNISSEN (*Kolloid-Z.*, 91 (1940) 311.

Phosphatides do not belong to macromolecular, but to association colloids. They have, as already stated (p. 188 and 262), been considered nevertheless in this Chapter of the book, as the characteristic charge properties do not depend on this discrimination, but on the properties of the ionised groups.

Phosphatides are extremely difficult to obtain in a pure state. The phosphatide preparations used must be considered as mixtures of pure phosphatides (ampholytes) and phosphatidic acid, the latter giving the phosphatide preparation the character of a "colloid of acidic nature". The purer the preparation the higher the reciprocal hexol-number, the higher also the *I.E.P.* and as shown in the figure the lower the reversal of charge concentrations for polyvalent cations (*e.g.*, Ca).
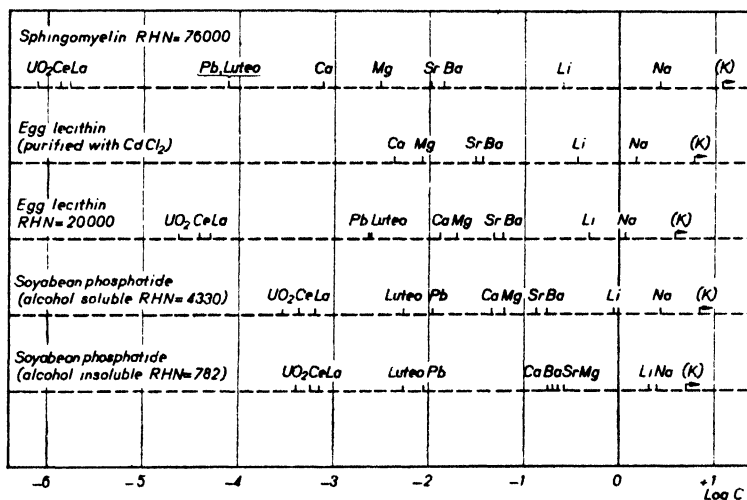
Fig. 20. Reversal of charge spectra for a number of phosphatide preparations with different reciprocal hexol numbers (RHN).
The number of cations in this figure is reduced to the minimum which suffices to bring out the phosphate colloid character of the phosphatides. This figure shows that the higher the reciprocal hexol number (*i.e.* the closer the phosphatide preparation approximates to pure phosphatide) the lower the reversal of charge concentrations come out (for the rest this displacement does not make its appearance with Na and K) (see further text).

A close dependence on the *R.H.N.* (as a measure of the apparent equivalent weight, see § 1 f, p. 273) exists further in the absolute values of the reversal of charge concentrations with large organic cations (see Fig. 24, 26 and 27 on p. 301, 303 and 304) and in the extent of the antagonism CaCl₂—NaCl (p. 312, § 5 b).

m. Final remarks

In § 1 e we discussed on the basis of Fig. 5 (p. 271), indications for the occurrence of a second characteristic charge element, which we have investigated in detail in this § 2. The results obtained induce us to reconsider the statements made in § 1 e and even those of § 1 d, dealing with the correlation between flocculability and equivalent weight.

It can now be understood why in § 1 d (p. 269) and § 1 e the sulphate colloids showed more or less the same behaviour as the phosphate and carboxyl colloids. We

compared these colloid types as regards their flocculability with 6, 5, 4 and 3 valent complex cations and with monoatomic lower valent cations (Ca and Na).

This choice of cations seems afterwards not to have been an adequate one. For as we saw in § 2 i (p. 291) complex ions, by their inability to polarise the ionised groups, behave in a different way from the small monoatomic cations which have this property.

It is even by this choice of 6, 5, 4, and 3 valent complex cations, that in § 1 d the sulphate colloids differ only quantitatively from the phosphate and carboxyl colloids. For in this case the fundamental difference in polarisability between sulphate groups on the one hand and phosphate and carboxyl groups on the other has no, or practically no, influence on the results.

If it were possible to use in Table 2 (p. 270) only ideal 6, 5, 4, and 3 valent mono-atomic cations (La and Ce being the only ideal cases of 3 valent cations, Al and Th already behaving as large complex ions, see p. 291, and the still higher valent cations not existing at all) then the sulphate colloids would behave totally otherwise than the phosphate and carboxyl colloids (horizontal rows consisting presumably only in minus signs).

For increase in valency of small monoatomic cations may have as a result in sulphate colloids, that reversal of charge is shifted to higher concentrations and flocculation becomes more difficult.

Indeed carrageen flocculates with $Co(NH_3)_6Cl_3$ (complex cation), but not with $La(NO_3)_3$ (monoatomic cation).

Before closing this section, we will consider the cation spectra of particles of SiO_2 and TiO_2 suspended in water. See Fig. 21.
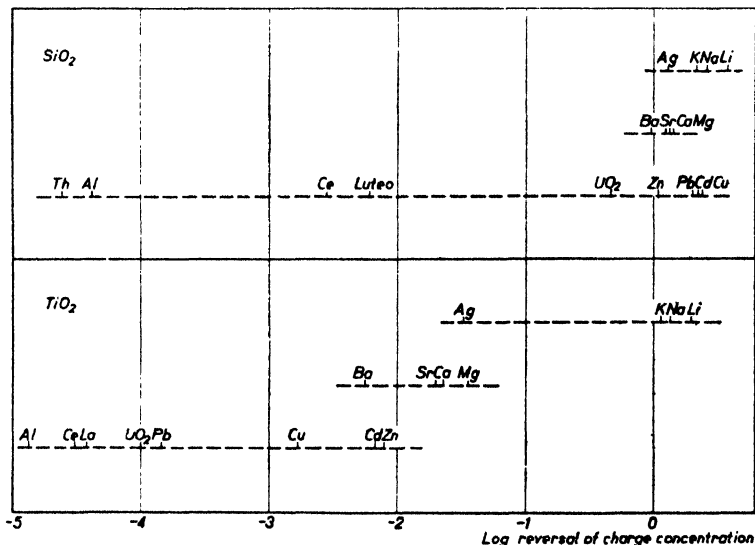
Fig. 21.   Reversal of charge spectra for SiO_2 and TiO_2 particles.
The reversal of charge spectrum of SiO_2 resembles fairly much that of sulphate colloids (p. 285, Fig. 15), that of TiO_2 the reversal of charge spectrum of carboxyl colloids (p. 284, Fig. 14).

By inspection of these ion spectra, it appears that the sequences of the alkali and alkaline earth cations are those also found in carboxyl and sulphate colloids. However as regards the relative positions of the cations of the B subgroups and of the A subgroups a marked difference exists in both spectra.

The ion spectrum of SiO_2 resembles nearly that of a sulphate colloid, the ion spectrum of TiO_2 that of a carboxyl colloid. It follows from the ion sequences that SiO_2 is less, TiO_2 is more polarisable than water.

We have to consider the origin of the negative charge as follows: SiO_2 forms at the surface of the particles silicic acid, which is dissociated, leaving the particle negatively charged. In the same way the negative charge of TiO_2 arises from the dissociation of titanic acid at the surface of the particles.

Because of the smaller radius of the central Si ion, the surrounding oxygen ions will be bound more strongly than by the larger central Ti-ion.

Therefore the oxygen in SiO_2 (and in silicic acid) will be less polarisable than in TiO_2 (and in titanic acid).

Thus already the larger polarisability of TiO_2— which is evident of the ion spectra — may be foreseen.

Data on refraction confirm this — VAN ARKEL and DE BOER [1] give for the refraction of the oxygen ion in SiO_2 3.56, in TiO_2 6.2 and in water 3.76. From it follows in fact that SiO_2 is less, TiO_2 is more polarisable than water.

Thus we have here two suspensions, neither organic lyophilic substances, nor colloids which might be dissolved molecularly or not, but only two inorganic crystalline powders possessing true electrolytic character at their surface. They have been measured with the same techniques and using the same criterion (reversal of charge point) as the biocolloids. The conclusions drawn from these results as regards polarisabilities are confirmed by refraction data. This all thus means an indirect support for the conclusion, that in measuring reversal of charge concentrations, measurements are performed on the *electrolyte character of the biocolloids*.

§ 3. SPECIFIC ION SEQUENCES IN THE REVERSAL OF CHARGE OF PROTEINS

a. Cation sequences in the reversal of charge of negative proteins

We have already repeatedly seen, that at the I.E.P. proteins are not really uncharged, but that at this pH the simultaneously present positive and negative charges only compensate one another exactly (p. 186, 193, 217).

If we wish therefore to investigate the behaviour of the ionised carboxyl groups or of the ionised positive groups in proteins it will be advisable to choose working pH-values which do not lie in the neighbourhood of the I.E.P. For only in that way can we minimise the complications of the presence of the ionised group of opposite sign (though we cannot wholly exclude such influences).

So with the aim of investigating the reversal of charge concentrations of cations in gelatin and casein pH values of about 9.7 were chosen. Such high pH values

---

[1] A. E. VAN ARKEL and J. H. DE BOER, *Chemische Bindung als elektrostatische Erscheinung*, Hirzel, Leipzig 1931. See p. 93.

have however the serious draw-back, that the choice of the cations available for such an investigation becomes limited to the alkali and alkaline earth cations.

For cations of higher valence or belonging to the B subgroups of the Periodic System are here no longer admissible because of their tendency to hydrolysis, which would bring about 1) a transformation of these cations into hydrolysis products, and 2) undesired pH changes.

L. TEUNISSEN-VAN ZYP[1] found the following sequence:

gelatin

Ca $<$ Mg $<$ Ba $<$ Sr $\cdots$ $\cdots$ Li $<$ Na $<$ K
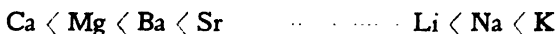
pH 9.67.

We observe that in gelatin the divalent cations cause reversal of charge at much lower concentrations than the alkali cations (the latter lie at concentrations higher than that of their saturated solution).

It is very interesting, that the sequence of the alkali cations is quite the reverse of that characteristic of the carboxyl colloids of § 2 d (see p. 284, Fig. 14), further that the sequence obtained with the alkaline earth cations is also different, a characteristic transition sequence[2] occurring. All this shows that the ionised carboxyl group in proteins is more polarisable than the same group in the carboxyl colloids of § 2 d. We have here once more a characteristic example of constitutional influences on the polarisability of the ionised groups, discussed already in § 2 k (p. 292).

We saw there that the sequence for oleates is:

Mg $<$ Ca $<$ Sr $<$ Ba $\cdot$ $\cdot$ Li $<$ Na $<$ K

and for the carboxyl colloids of § 2 d,

Ba $<$ Sr $<$ Ca $<$ Mg $\cdots$ $\cdot$ K $<$ Na $<$ Li.
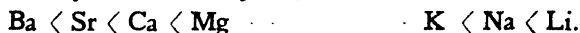
In these latter the polarisability is much decreased because of the presence of hydroxyl groups in the carbohydrate rings to which the carboxyl groups are attached. In proteins however practically all carboxyl groups stand at the end of aliphatic side-chains, and are separated by one (aspartic and hydroxyglutamic acid residues) or two (glutamic acid residue) CH_2 groups from carbon atoms attached to groups (peptide-, OH-) which may cause a decrease of polarisability of the carboxyl groups. Therefore it may be expected that the carboxyl groups in proteins stand much nearer to those in oleates than the carboxyl groups in the carboxyl colloids of § 2 d.

It is of course erroneous to speak of the behaviour of *the* carboxyl group of proteins, the reversal of charge sequences in reality informing us about the mean polarisabilities of the different categories of carboxyl groups present.

Now turning to casein, the reversal of charge concentrations found by L. TEUNISSEN-VAN ZYP[1] for casein are much the same as that found for gelatine:

casein

Ca $<$ Mg $<$ Sr $<$ Ba $\cdots$ $\cdots$ Li $<$ Na $<$ K
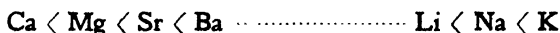
pH 9.74

Here once more the reversal of charge concentrations with the alkali cations lie at higher concentrations than those of the saturated solutions.

---

[1] L. TEUNISSEN-VAN ZIJP, *Thesis*, Leiden 1938 (Dutch).

[2] More variants of this transition series were observed with gelatin at lower pH values (e. g. at pH 6). They will not interest us here as complications may be involved when approaching the I.E.P., see p. 297.

The only difference with gelatin is the relative position of Sr and Ba in the transition series of the alkaline earth cations.

In casein, two kinds of negative ionised groups occur, here phosphate groups (attached to the hydroxyl of serine) being present next to carboxyl groups. As we should already expect for these phosphate groups, if exclusively present, a transition series for the alkaline earth cations and the sequence Li $<$ Na $<$ K (see p. 295, Fig. 20), it will not cause surprise at all that casein does not give a quite different ion spectrum as found in gelatin.

b. Anion sequences in the reversal of charge of positive proteins

In positive proteins different positively charged ionised groups may be present, as the following three diamino-mono-carboxyacids: arginine, lysine and histidine, may take part in building up the protein molecule. Thus guanidine, amino and imidazol groups may be present at the end of side chains as ionised groups. The simplest case is that of clupein, in which only arginine is present, then follows gelatin in which arginine and histidine take part and finally casein possessing all three kinds of basic amino acids. L. TEUNISSEN-VAN ZYP therefore chose these three proteins in a preliminary investigation on the reversal of charge concentrations of positively charged proteins. Though having observed that here the valency of ions is a factor of primary importance (the reversal of charge concentration increasing in the case of casein at pH 3.4 markedly in the order $K_4Fe(CN)_6$ $<$ $K_3Fe(CN)_6$ $<$ K_2SO_4 $<$ KCl), she restricted the investigation to a comparison of the alkali salts KCl, KBr, KJ, KNO$_3$ and KCNS [1].

These results are given in Fig. 22
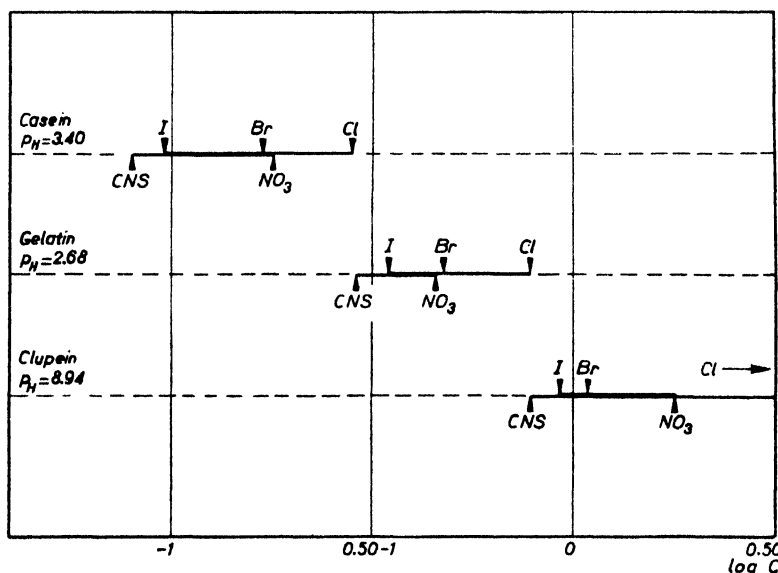
Fig. 22. Reversal of charge spectra (monovalent anions) of some positively charged proteins. The sequence I $<$ Br $<$ Cl occurs for the monatomic anions; for the polyatomic CNS $<$ NO$_3$.

---

[1] KF being the salt of the weak acid HF cannot be used, as it changes the pH very considerably.

from which it may be seen that in all three proteins nearly the same sequence is obtained.

$$CNS \langle J \langle Br \quad \langle NO_3 \langle Cl \qquad \text{casein (pH 3.4) and clupein (pH 8.9)}$$
$$CNS \langle J \langle NO_3 \langle Br \quad \langle Cl \qquad \text{gelatin (pH 2.7)}$$
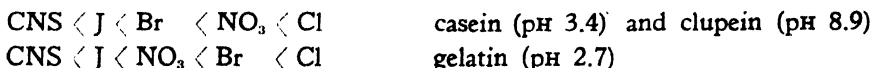
only differences in the order of Br and NO_3 occurring.

But from a theoretical point of view it has no sense to compare monoatomic anions with polyatomic anions. We have thus to compare the halogen ions and find then for all three proteins the same sequences

$$J \langle Br \langle Cl$$
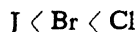
This sequence is liable to two different explanations as regards the polarising power of the positive ionised groups of proteins.

1. These groups, being too large, have practically no polarising power. For then the fixation of the anions will be the more easy, the greater their radius, that is the less the hydration of the halogen ions.

2. If the positive ionised groups of proteins do exert a noticeable polarising power, then the sequence following from 1. *viz.* $J \langle Br \langle Cl$ will not be altered.

For from this supposed extra polarisation effect by itself the sequence $J \langle Br \langle Cl$ would also follow, as the polarisilibity of the halogen ions decreases in the sequence $J—Br—Cl$.

Anticipating here results to be discussed in § 4, in which it is found that substituted ammonium cations do not produce noticeable polarisation of the negative ionised groups of colloids of acidic nature, it seems that explanation 1 comes nearest to the truth (see p. 309).

§ 4.   REVERSAL OF CHARGE WITH ORGANIC IONS

a. Introduction

Monovalent inorganic cations generally have high reversal of charge concentrations in negative colloids of acidic nature (seldom lying lower than 1 N).

With organic cations they may lie according to circumstances lower or even very much lower, as first was met with in investigating coacervation phenomena of colloids with dyes [1]. A first systematic investigation on the action of organic cations was made by BUNGENBERG DE JONG and WAKKIE [2] who studied the reversal of charge of four phosphate, two carboxyl and two sulphate colloids with four randomly chosen organic cations on the one hand and with Li, Na and K on the other. Fig. 23 gives as an example the interpolation graph for an egg lecithin sol, showing the great spread in the reversal of charge concentrations, some alkaloid cations being even more active than the divalent Ca ion. Fig. 24 and 25 give a survey of the results obtained.

The survey shows that though the sequence of the alkali cations in phosphate

---

[1] H. G. BUNGENBERG DE JONG and J. LENS, *Bioch. Z.*, 254 (1932) 15.

Unpublished investigations show even, that among basic dyes examples may be found, which because of their very low real reversal of charge concentrations, might possibly be used in a similar way as discussed for hexol nitrate in § 1 b (p. 262), for obtaining similar "reciprocal dye numbers". Such dyes should however be available in a well defined and absolutely pure state.

[2] H. G. BUNGENBERG DE JONG and J. G. WAKKIE, *Bioch. Z.*, 297 (1938) 70 and 221.

colloids is the reverse of that in carboxyl and sulphate colloids, no such reverse is met with in the sequence of the four organic cations, this being the same in all colloids:

$$\text{quinine} < \text{strychine} < \text{procaine} < \text{guanidine}.$$
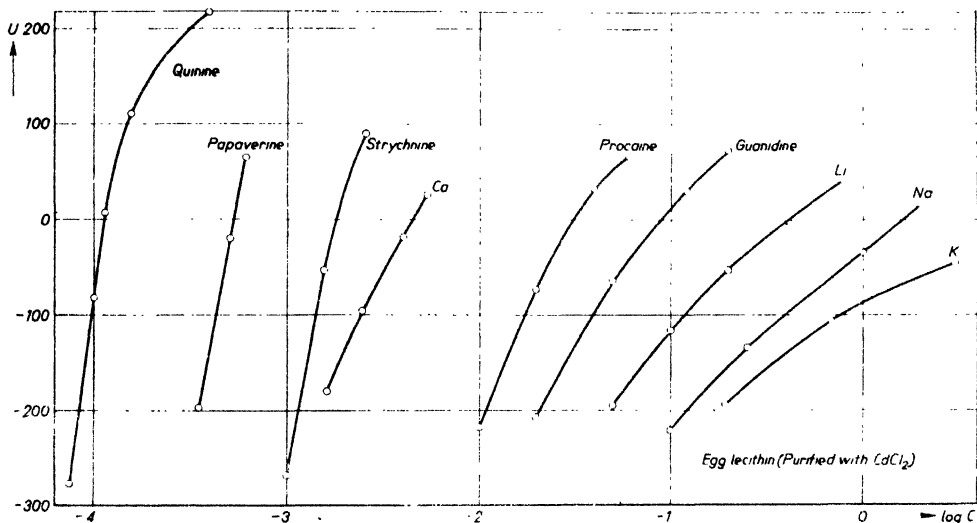
Fig. 23. Reversal of charge of egg lecithin (purified via the $CdCl_2$ compound) with some alkaloid cations and with some inorganic cations for comparison (see text).

As the reversal of the sequence of the alkali cations is the result of polarisation effects (see p. 287, § 2 f) we may conclude that the organic cations used have no polarisation power, which could of course be expected from their relatively large volumes.

Fig. 24. Reversal of charge spectra of phosphate colloids with hydrochlorides of quinine, strychnine, procaine, guanidine; LiCl, NaCl and KCl. The sequence of the organic cations is from left to right: quinine — strychnine — procaine — guadinine, except with alcohol soluble soya bean phosphatide where the reversal of charge points of procaine and guanidine are interchanged (the sequence of the U—log C curves is however the normal one up to very close to the reversal of charge point).
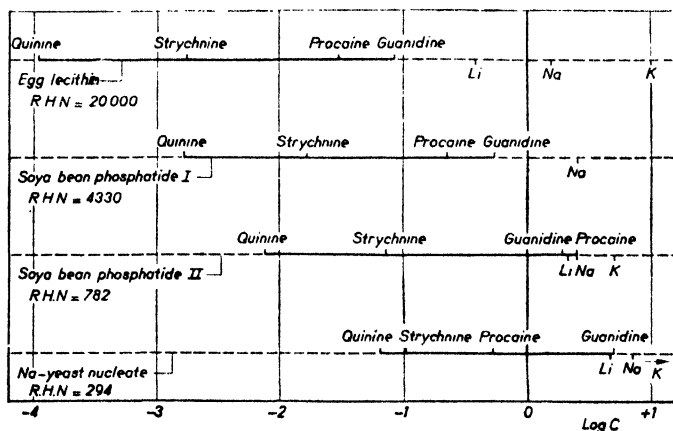
Seeking an answer to the question, why the above sequence is obtained, we observe that guanidine, having the simplest structure shows the highest reversal of charge concentration, and that proceeding in the sequence to the left the organic chemical structure of the cations becomes more complicated.
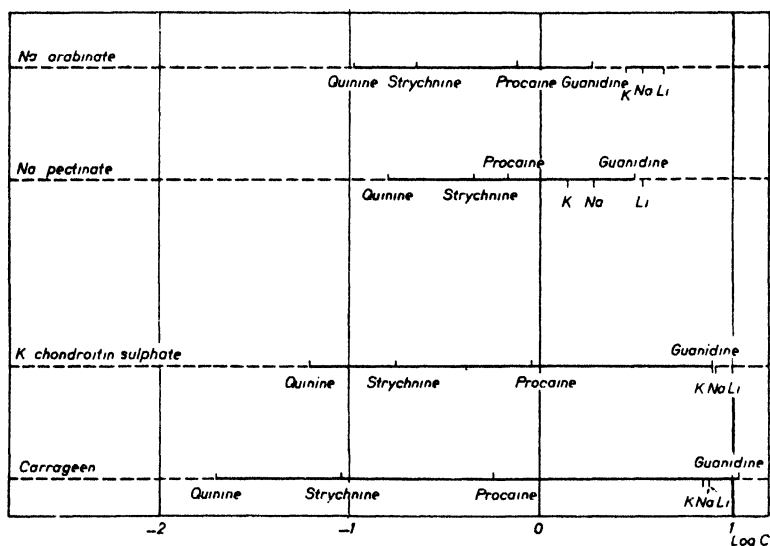
Fig. 25. Reversal of charge spectra of two carboxyl and two sulphate colloids with hydrochlorides of quinine, strychnine, procaine, guanidine; LiCl, NaCl and KCl.
Although the sequence of the alkali cations is now the opposite of that in Fig. 24, the sequence of the four organic cations has remained the same.

Thus we may come to the conclusion, that with organic cations increase of complexity of structure lowers the reversal of charge concentration *i.e.*, increases the affinity of the cation for the ionised groups of colloids.

In general analogous investigations, using a greater number of alkaloid cations will not give much principal insight into the relations between structure of the ion and relative affinity for the ionised groups[1].

For this aim one should begin with ions belonging to simple classes of substances, such as amines or fatty acids, which allow for a systematic change in the number and relative positions of the carbon atoms.

---

[1] Minor changes in structure, such as methylation or acetylation of hydroxyl groups may for instance be studied in a series of alkaloids related to morphine.
So H. G. BUNGENBERG DE JONG and C. VAN DER MEER, Proc. Koninkl. Nederland. Akad. Weten-schap., Amsterdam, 45 (1942), 593 found for a soya bean phosphatide (soluble in alcohol) the following sequence of from left to right increasing reversal of charge concentrations;

apomorphine < thebaine < heroine < ethylmorphine < morphine < codeine

from wich it seems probabie that acetylation or alkylation increases affinity and further that alcoholic hydroxyl and phenolic hydroxyl have here a reverse influence, the first decreasing and the latter increasing affinity.

This problem was attacked by L. TEUNISSEN-VAN ZYP [1].

As the results obtained are as yet only published in Dutch [1] and therefore not generally accessible, a survey will be given in the following subsections b—d.

b. Reversal of charge sequences of substituted ammonium cations in sulphate, carboxyl and phosphate colloids

Earlier investigations have shown, that the ammonium ion in general resembles the K ion or stands between Na and K as regards its place in reversal of charge sequences (see p. 289, Fig. 17). Now we can derive organic cations in two ways from the ammonium cation:

a. by replacing one H in NH_4^+ by alkyl groups of increasing chain length,

b. by replacing one, two, three and four H in NH_4^+ by the same alkyl groups, and ask what will be the result on the reversal of charge concentration.

For this investigation a sulphate colloid (Na agar), a carboxyl colloid (Na pectinate) and three phosphate colloids (Na yeast nucleate, purified egg lecithin, and a soya bean phosphatide fraction soluble in alcohol) were used.

Unfortunately in many cases (all amines in agar; very many in nucleate), the reversal of charge concentrations lay at such high concentration that the reversal of charge concentration could not be reached.

The sequence of the curves is however quite the same as in colloids in which all or a larger number of reversal of charge points could actually be reached. The ion spectra of the latter colloids, viz., egg lecithin and soya bean phosphatide (phosphate colloids) and Na pectinate (carboxyl colloid) are given in Fig. 26, 27 and 28.
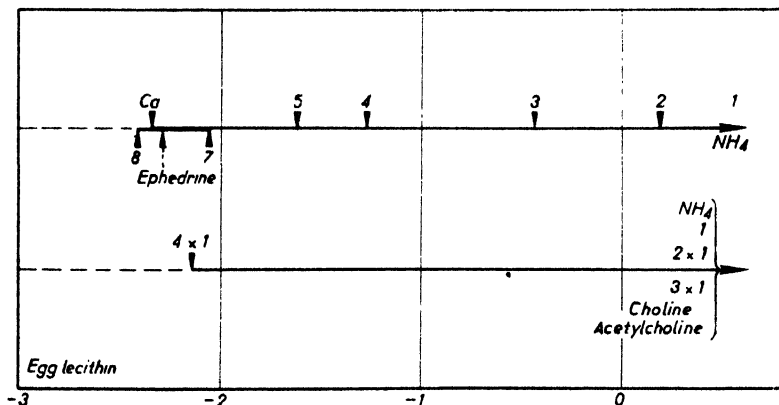
Fig. 26. Reversal of charge spectra of egg lecithin (R.H.N > 20000) with substituted ammonium cations.
Upper series: hydrochlorides of NH_3 and primary amines:
1 = methylamine (reversal of charge concentration like that of NH_4^+ so high that it cannot be reached) 2, 3, 4, 5 = ethyl, n propyl, n butyl and isoamylamine, 7 and 8 = benzylamine and phenylethylamine. In addition ephedrine and Ca'' are also included here.
Lower series: 1 = methylamine; 2 × 1 = dimethylamine; 3 × 1 = trimethylamine, all three. as indeed also choline and acetylcholine, with unreachably high reversal of charge concentrations, In contrast with this 4 × 1 = tetramethylammonium chloride has a fairly low reversal of charge concentration.

---

[1] L. TEUNISSEN-VAN ZIJP, *Thesis*, Leiden 1938.
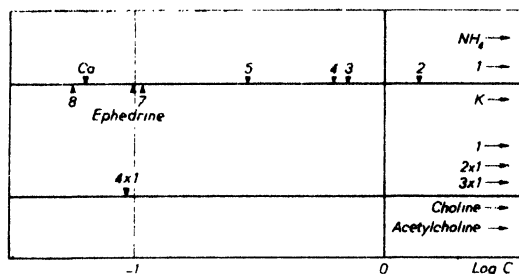
Fig. 27. Reversal of charge spectrum of alcohol soluble soya bean phosphatide (R.H.N = 3800) with substituted ammonium cations.
Significance of the numbers at the reversal of charge points as in Fig. 26. The sequence of the various cations is the same as with egg lecithin, although (apart from 2 and 3) they lie at higher absolute concentrations (as is also the case in Fig. 20, p. 295 for inorganic cations and in Fig. 24, p. 301, for alkaloid cations).

Fig. 28. Reversal of charge spectrum of Na pectinate (carboxyl colloid) with substituted ammonium cations. Significance of the numbers at the reversal of charge points as in Fig. 26.
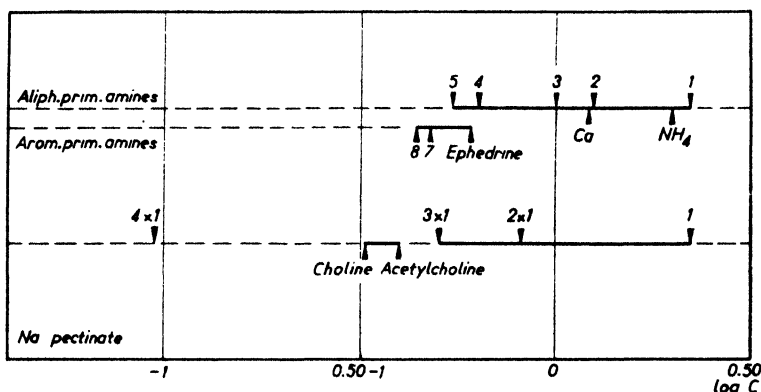
In contrast with the phosphatides the reversal of charge is quite reachable in this case with NH_4, 1 ,2 × 1 ,3 × 1 , acetylcholine and choline.

Further also all reversal of charge points could be reached with SiO_2 particles (without colloids). As we have already seen that the ionogenic surface of SiO_2 very much resembles the sulphate colloids in polarisability (see p. 296 § 2 m), the results obtained on SiO_2 have also been given in Fig. 29 to serve as a substitute for the cation sequences in sulphate colloids.

Fig. 29. Reversal of charge spectrum of SiO_2 particles (substitute for sulphate colloid) with substituted ammonium cations. Significance of the numbers at the reversal of charge points as in Fig. 26.
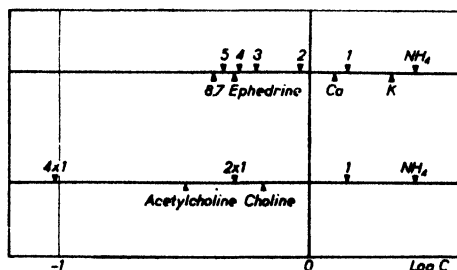
In these figures the individual amines have been denoted by the number of carbon atoms in the carbon chain(s). Thus the upper row(s) represent hydrochlorides

of aliphatic primary amines: methyl- (1); ethyl- (2); n. propyl- (3); n. butyl- (4) and isoamylamine (5) and two aromatic primary amines: benzylamine (7) and β-phenyl-ethylamine (8).

The lower row represents methyl- (1), dimethyl- (2 × 1), trimethylamine (3 × 1) and tetramethylammonium (4 × 1) hydrochlorides.

On both rows some additional reversal of charge concentrations are given, which from a physiological point of view might have a certain interest[1].

The results obtained, may be summarised as follows:

1. The introduction of the first CH_3-group in NH_4^+ shows no general rule, it either lowering or increasing somewhat the reversal of charge concentrations.

2. Lengthening of the carbon chain in primary amines lowers the reversal of charge concentration:

   isoamylamine ⟨ butylamine ⟨ dimethylamine ⟨ methylamine

and

   b-phenylethylamine ⟨ benzylamine.

3. Introduction of more CH_3-groups in methylamine lowers the reversal of charge concentrations:

   tetramethylammonium ⟨ trimethylamine ⟨ dimethylamine ⟨ methylamine.

The points 2 and 3 show in the first place that these relatively large organic cations do not exert a noticeable polarising power, for the sequences obtained are the same for sulphate, carboxyl and phosphate colloids (in distinction to the smaller inorganic alkali cations, which show in phosphate colloids the reverse sequence of that in carboxyl and sulphate colloids).

The lowering of the reversal of charge concentration summarised in point 3 is understandable without further assumptions from the increase in volume of the ion by introducing more and more methyl groups around the central N atom. Contrary to this it is not directly obvious why the lengthening of the carbon chain in primary amines (point 2) has the same effect. We shall return to this question after having discussed analogous effects in organic anions (see p. 309, § 4 d).

The following points are still of interest:
1. The very great influence in lowering the reversal of charge concentration by introducing a fourth methyl group into trimethylamine, is only present in phosphatides, (compare Fig. 26 and 27 in which 1, 2 × 1, 3 × 1 lie at very high concentration, 4 × 1 is however found at a relatively low concentration.

---

[1] It is interesting that adrenaline and acetylcholine show in certain physiological conditions Ca- and K-like actions.

Instead of the very readily oxidisable adrenaline, ephedrine was taken. Both adrenaline and ephedrine are in a certain sense to be considered as derivatives of b-phenylethylamine and therefore the reversal of charge concentration of ephedrine is set out on the upper row(s) of the figures.

On the lower row choline and acetylcholine are set out, both being derivates of trimethyl-ethyl-ammonium chloride, thus being related to 4 × 1.

The experimental results show that a simultaneous resemblance of ephedrine and Ca on the one side and of acetylcholine and K on the other side is present neither in sulphate colloids (agar) nor its substitute (SiO_2), nor in carboxyl colloids (pectinate) nor in the phosphate colloid nucleate but only in phosphatides.

2. Introduction of alcoholic hydroxyl seems to exert an increase of reversal of charge concentration. For choline lies always at a higher concentration than tetramethyl ammonium, this difference being extraordinarily great in phosphatides (see Fig. 26 and 27). Now the choline ion $(CH_3)_3$. N . C_2H_5OH, contains one C atom more than the tetramethyl-ammonium ion and for this reason it should have a lower reversal of charge concentration. The increase actually found is thus an indication for a powerful effect of the OH group in the reverse direction.

The reversal of charge concentration of ephedrine lies at higher concentration than of β-phenyl-ethylamine. This is also an indication for a marked increasing action of alcoholic hydroxyl. For ephedrine, compared with β-phenyl ethylamine, contains not only a hydroxyl group more, but also two carbon atoms more, which latter would lower the reversal of charge concentration.

c. Reversal of charge sequences of organic anions in positive proteins

With three positively charged proteins, the reversal of charge sequence was investigated for three sodium alkyl-ester-sulphates. The results (see Fig. 30) showed the same sequence for all three proteins:

isoamyl sulphate $<$ isobutyl sulphate $<$ propyl sulphate.

Thus, as was the case with the cations of the primary amines in negative colloids, (§ 4 b) here also increasing the number of carbon atoms of an organic anion decreases the reversal of charge concentration in positive colloids.
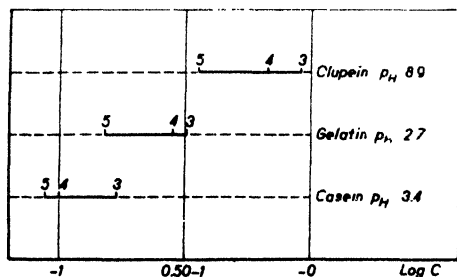
Fig. 30. Reversal of charge spectrum of some positively charged proteins with propyl sulphate (3); isobutyl sulphate (4) and isoamyl sulphate (5).

Now the ester sulphates were chosen here because of the corresponding acids being strong acids. This makes it possible to investigate the influence of these anions at pH values in which casein and gelatin are sufficiently positively charged.

The vast amount of organic acids, being weak or relatively weak carboxyl acids — the alkali salts of which would constitute an ideal arsenal for investigating the relation between constitution of organic anions — cannot be used with gelatin and casein.

To exclude all hydrolysis for these salts a much higher pH, lying even on the alkaline side of pH 7 should be chosen, excluding thus all so-called "acid" proteins, that is, proteins having their I.E.P. below pH 7.

Of the "basic proteins" only protamins (having their I.E.P. at very high pH values) come into consideration. For this investigation clupein sulphate was chosen, the I.E.P. of which was determined electrophoretically to be pH = 13.45 (see Fig. 45 on p. 325). The reversal of charge concentrations of a considerable number of organic anions were determined at a pH 8.94 of the blanks.

Fig. 31 shows the results obtained with the sodium salts of eleven aliphatic monocarboxylic acids (upper rows) ranging from formate to stearate, each individual being indicated in the figure by the total number of carbon atoms.

Just below this upper row are given the results obtained with the sodium salts of three aromatic acids, —benzoic, phenylacetic and phenylproionic acid — each characterised in the figure by the total number of carbon atoms, thus 7, 8 and 9.

We see that the sequence of anions of the fatty acids is:

Stearic < myristic < lauric < capric < caprylic < caproic < valeric < n. butyric <
  (18)   (14)   (12)   (10)   (8)   (6)   (5)   (4)

propionic < acetic < formic
  (3)   (2)   (1)

and of the aromatic monocarboxy acids:

phenylpropionic < phenylacetic < benzoic
    (9)   (8)   (7)
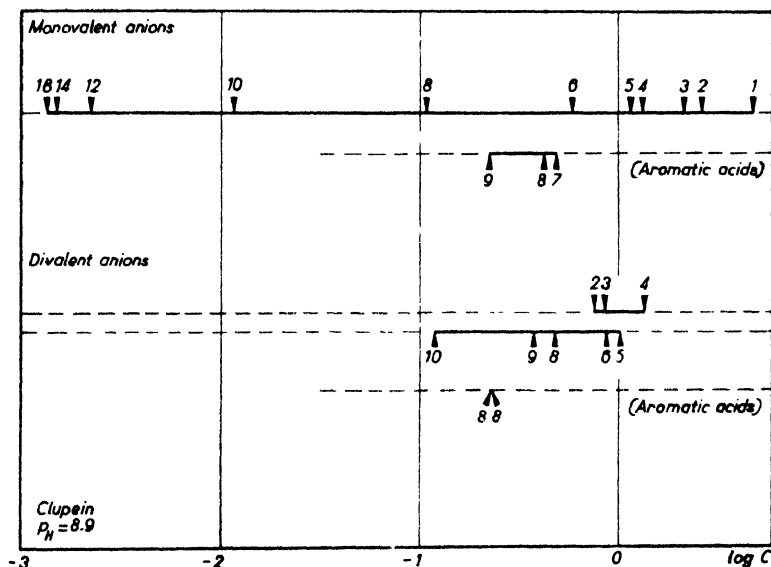
Fig. 31. Reversal of charge spectrum of positively charged clupein (pH ≈ 8.94) with mono and divalent organic anions.
Top horizontal rows:
1, 2, 3, 4, 5, 6, 8, 10, 12, 14, 18 = anions of unbranched fatty acids, named according to the number of C atoms.
7 = benzoate, 8 = phenylacetate, 9 = phenylproprionate.
Next two rows:
2, 3, 4, 5, 6, 8, 9, 10 = anions of the oxalic acid series (oxalic acid, malonic acid, etc., to sebacic acid).
Lowest level:
8 and 8: terephthalate and orthophthalate.
While the successive members of the acetic acid series each time show reversal of charge at a lower concentration the more the C chain of the ion increases in length, with the oxalic acid series a rise of the reversal of charge concentration first occurs (2 ⟶ 3 ⟶ 4) and then afterwards a fall (10 ⟵ 9 ⟵ 8 ⟵ 6 ⟵ 5 ⟵ 4).

Thus we meet here with the same rule as in the case of the alkylester sulphates — but now on much more extensive experimental material — that lengthening the carbon chain of an organic anion decreases the reversal of charge concentration.

The decrease in reversal of charge concentrations is very considerable on increasing the number of carbon atoms above a certain chain length (of about 6 C atoms), but seems to come at an end at about a chain length of 12 C atoms. The crowding together of 12, 14 and 18 (and also of 16) finds however its explanation in the fact that here the reversal of charge concentrations are no longer true ones. For already stearate lies very near to the equivalent amount of the clupein present and it was explained in § 1b (p. 263) and 2a (p. 276), why in such a case the reversal of charge concentration becomes fictitious and represents nearly the amount of ions fixed to the ionised groups of the amount of colloid present.

Very interesting results were obtained with the sodium salts of the dicarboxylic acids. They form an irregular series, which can however be written in the way as given in Fig. 31 as two regular sequences connected by the succinate ion:

<div align="center">

oxalic $<$ malonic $<$ succinic

(2)   (3)   (4)

</div>

<div align="center">

sebacic $<$ azelaic $<$ suberic $<$ adipic $<$ glutaric $<$ succinic

(10)   (9)   (8)   (6)   (5)   (4)

</div>
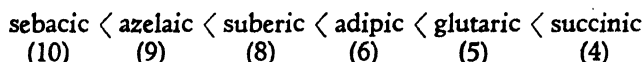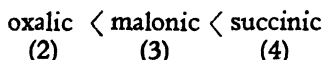
Thus from succinate onwards, lengthening of the carbon chain between the two ionised carboxyl groups lowers the reversal of charge concentration, but starting from the oxalate ion the first two extra CH_2 groups have the reverse effect.

The following explanation of the occurrence of this curious transition sequence starts from the idea, that we have here no real exception to the rule frequently found that lengthening of the carbon chain lowers the reversal of charge concentration. It is thought that this influence is present in the whole series of dicarboxylic acids, in the first terms of the series it is only overcompensated by another influence acting in the opposite direction.

Introducing CH_2 groups in oxalic acid not only builds up a carbon chain but separates also the two negative charges, which originally were close together.

The oxalate ion stands not far from a divalent anion in which both negative charges coincide (e.q. sulphate ion). In accordance with it, it has a much lower reversal of charge concentration than the monovalent formate ion.

Now introducing an increasing number of CH_2 groups, the two negative charges become more and more separated, which will cause in itself an increase of the reversal of charge concentration. For by separating the two ionised groups, the divalent ion assumes the morphology of two monovalent ions bound together (e.g., the succinate ion $= 2 \times$ the acetate ion).

This increasing influence on the reversal of charge concentration will be of course greatest on introducing the first CH_2 groups. In introducing more this influence will be much smaller and will no longer overcompensate the general lowering influence of the increase in length of the carbon chain.

Further results obtained with clupein may be shortly mentioned, which throw some light on the connection between structure of the anion and its affinity for the ionogenic groups.

1. The sequence:

<div align="center">

phenylpropiolic acid $<$ cinnamic acid $<$ phenylpropionic acid

</div>

shows the increase in affinity by a triple and a double carbon — carbon bound.

2. The sequence:

> salicylic acid < p. oxybenzoic acid < benzoic acid

seems to reveal an increase in affinity by the introduction of a phenol hydroxyl [1]. Still this is uncertain as in the medium of pH 9 a beginning of phenolate formation may exert its effect.

3. Halogen substitution in aliphatic and aromatic acids increases affinity as shown by the sequence

> trichloracetic < dichloracetic < monochloracetic < acetic-acid.

and

> p jodobenzoic- < p chlorbenzoic- < benzoic-acid.

4. No general rule was found for the relative positions (o. m. p.) of the substituents.

5. Replacing a benzene group by a naphthalene group increases the affinity considerably.

> α and β naphthalene sulphonic- < benzene sulphonic-acid

d. Discussion

Summarizing the findings surveyed in § 4 b and § 4 c we may generally say that polarisation effects do not play a rôle in the fixation of organic ions on the ionised groups of colloids. This is evident for organic cations from the similar behaviour of phosphate, carboxyl and sulphate colloids towards amine hydrochlorides discussed in § 4 b.

One might ask if the positively charged ionised groups in proteins ($- NH_3^+$; imidazol group of histidin and the guanidinium group in arginin) may still exert a polarising action in fixing anions. This question, already put as regards the fixation of inorganic ions in § 3 b (p. 299), could not be settled there, but we may now conclude with high probability from the absence of polarisation effects of substituted ammonium cations in § 4 b, further from the normally very high reversal of charge concentration of guanidine hydrochloride for all phosphate, carboxyl and sulphate colloids (§ 4 a), that the above-named positive groups do not exert polarising actions in the fixation of anions.

The only experimental finding, that at first sight suggests the presence of polarisation effects, namely the irregular series of the dicarboxylic acid anions — constituting a typical transition sequence —, could be explained on a quite different basis (§ 4 c).

In general aspects the influence of structure on reversal of charge concentrations of organic cations and anions is quite the same, the outstanding fact being that lengthening of the carbon chain decreases the reversal of charge concentration. This is so much the more remarkable because of the hydrocarbon chain, being a structural part carrying no free charge, seems to play such a prominent part in the undoubtedly electrical mechanism of ion fixation on the ionised groups of the colloid. Indeed we cannot accept such a rôle and we must therefore discuss other possibilities of explaining the influence of chain length on reversal of charge concentration.

---

[1] It seems dangerous to conclude from the observed sequence

> tartrate < malate < succinate

that introduction of alcoholic hydroxyl increases affinity. For the succinate ion stands just on the "bend" of the transition sequence of the dicarboxylic acids, which may complicate matters. Such an influence which would be the reverse of that found in organic cations (small print on p. 302) seems also not very probable from the discussion which will follow below.

At first it seems that a special affinity of the carbon chain for certain non-ionised groups of the colloid may be responsible for this, which adds itself to the electrical affinity proper between the two charged parts (ionised groups in colloid and organic ion) involved in the salt formation.

This seems quite possible in the case of proteins, for the macromolecule here has a certain fraction of side chains of exclusively hydrocarbon character, which latter may serve as "fixation-spots" for the hydrocarbon chains of organic ions.

However difficulties arise for the generality of the above explanation if we consider that the influence of chain length is found just as much in pectinates and agar, the latter constituting macromolecules of the polymeric carbohydrate type, in which no such "fixation-spots" of exclusively hydrocarbon nature are present. Every carbon atom is here linked to oxygen somehow (hydroxyl groups, ether links) and the macromolecule will therefore have on its whole "surface" a pronounced hydrophilic character.

We are thus induced to seek the main explanation elsewhere — though we will not deny that in proteins the above suggested mechanism might play a rôle in addition.

Thus it is proposed that in the fixation equilibrium — which is characterised by the true reversal of charge concentration — the tendency of the organic ions to escape from the aqueous medium in which they are embedded plays an important part. In a homologous series of organic ions e.g., of amines or of fatty acids, this tendency is least in the first terms and increases in lengthening the hydrocarbon chain — manifesting itself in decreasing solubilities or (and) increasing tendency to form aggregates in solution (in the fatty acid series a transition is found from mono ionic solutions to sols of association colloids).

It is assumed that the greater this "escaping tendency", the more easy the fixation of the organic ions on the ionised groups will be, that is the reversal of charge concentration will decrease with increasing carbon chain length.

We might express this also in the following manner: in general the reversal of charge concentration will diminish if such changes in an organic ion are brought about, which render it as a whole more hydrophobic (lengthening of carbon chain, replacing H by halogen atoms, replacing a benzene nucleus by a napthale.ie nucleus and so on), whereas introduction of groups which render it as a whole more hydrophilic may have the reverse effect.

We are aware that the explanation given is still very schematic, also not enough experimental material is at hand to check adequately the predictions based on this view. As to alcoholic hydroxyl see note 1 on p. 302, small print on p. 306 and note on p. 309.

Certain exceptions may also be foreseen. Thus phenolic hydroxyl seems to lower reversal of charge concentrations in certain colloids [1], though introduction of it increases the hydrophilic character of the anion. But in the case of proteins this is quite conceivable as phenolhydroxyl has here a special affinity for certain groups within the protein molecule (see small print on p. 255).

The supposed mechanism might also give an explanation, why at sufficiently long chain length or at sufficient complexity of structure of the organic ion not only the reversal of charge concentration becomes small but also why then flocculation of the colloid is a frequent phenomenon.

For the "escaping tendency" of an organic ion will be most satisfied if its hydrophobic chains or carbon structures may serve also in uniting the separate colloid particles on which they are attached by their ionised spots.

---

[1] In phosphatides: see note on p. 302; in clupein: salicylic acid $<$ benzoic acid, see small print in § 4 c on p. 309.

§ 5.   REVERSAL OF CHARGE WITH SALT MIXTURES

a. Introduction. Additivity and antagonism

We first describe the behaviour of phosphatide sols towards mixtures of salts having the same anion in common. If the reversal of charge concentration is determined with $CaCl_2$, first without an added second salt and then in the presence of some constant, not too high, NaCl concentrations, it is invariably found that in the latter cases the $CaCl_2$ concentration needed to reach the reversal of charge point is higher than in the blank experiment.

At higher NaCl concentrations the $CaCl_2$ concentration reaches a maximum and at still higher concentrations the $CaCl_2$ concentration decreases first slowly, than rapidly and of course becomes zero if the NaCl concentration chosen is the reversal of charge concentration of NaCl itself.

In Fig. 32 a (upper curve) the results are given for a soya bean phosphatide fraction (alcohol-soluble)[1]. In this figure the reversal of charge concentration for each of the two salts separately is taken as 100% ($CaCl_2 = 0.045$ N; NaCl $= 2.8$ N). We see that in this case at a definite NaCl concentration the $CaCl_2$ concentration needed for reaching the reversal of charge point is more than twice as high (224%) as in the blank.

Fig. 32. Reversal of charge of alcohol soluble soya bean phosphatide with mixtures of LiCl + NaCl, $CaCl_2$ + NaCl and $Co(NH_3)_6Cl_3$ + NaCl.
Ordinates: concentrations of $Co(NH_3)_6Cl_3$ or $CaCl_2$ or LiCl in the salt mixture expressed in % of the reversal of charge concentrations of these salts in the absence of NaCl.
Abscissae: concentration of NaCl in the salt mixture expressed in % of the reversal of charge concentration of NaCl in the absence of another salt.
The curves must therefore always begin at 100% on the ordinate axis and end at 100% on the abscissa axis. Only slight deviations from additive behaviour (dotted straight line) are present in the combinations LiCl + NaCl. In the other two combinations antagonism occurs, which is considerably stronger in the combination $Co(NH_3)_6Cl_3$ + NaCl than in the combination $CaCl_2$ + NaCl.
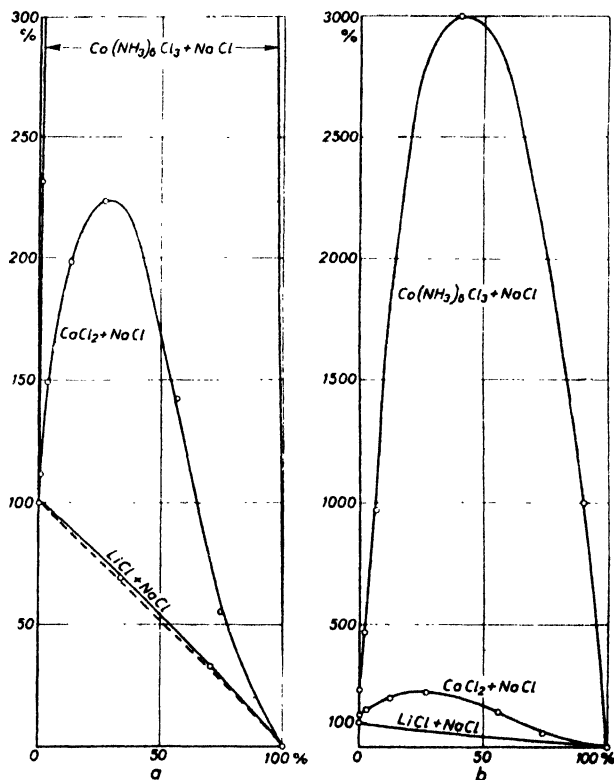
[1] H. G. BUNGENBERG DE JONG, H. L. BOOIJ and J. G. WAKKIE, *Kolloid-Beihefte*, 44 (1936) 254.

Now turning to other salt combinations, it is found that the deviations from additivity manifest themselves in principle always in the same way, if we only set out on the ordinate axis that salt which has the smallest reversal of charge concentration of the two.

Only the amount of deviation from additivity varies considerably according to the salt pair chosen.

Two further salt combinations determined on the same sol are also represented in the figure namely LiCl -| NaCl and $Co(NH_3)_6Cl_3$ + NaCl.

The first of the two combinations (see Fig. 32 a) shows only slight positive deviations from additivity, the second (Fig. 32b upper curve) much greater deviations than the combination $CaCl_2$ |- NaCl, here the maximum of the combination curve lying at a $Co(NH_3)_6Cl_3$ concentration which is about 30 times (3 000%) as high as the blank. For representing the curve of this combination the Fig. 32a has too large a scale in the ordinate direction, it had to be redrawn in Fig. 32b on a ten-fold smaller scale, to allow a convenient simultaneous representation of the combinations $CaCl_2$ + NaCl and $Co(NH_3)_6Cl_3$ |- NaCl.

Under *antogonism* proper we shall understand that the curve in a combination figure starts at 100% on the ordinate axis in a mounting direction (the two last named combinations).

b. Importance of the quotient of the reversal of charge concentrations of both members in a salt combination

We have seen above that in the salt combination LiCl + NaCl only small deviations from additivity occur (see Fig. 32). This is also the case with the combinations $CaCl_2$ -|- $MgCl_2$, (which shows practically additivity) and $La(NO_3)_3$ + $Co(NH_3)_6Cl_3$ (see Fig. 33 [1]).

In the combination $MgCl_2$ -|- NaCl a distinct antagonism occurs, very much comparable to that in the combination $CaCl_2$ + NaCl [2].
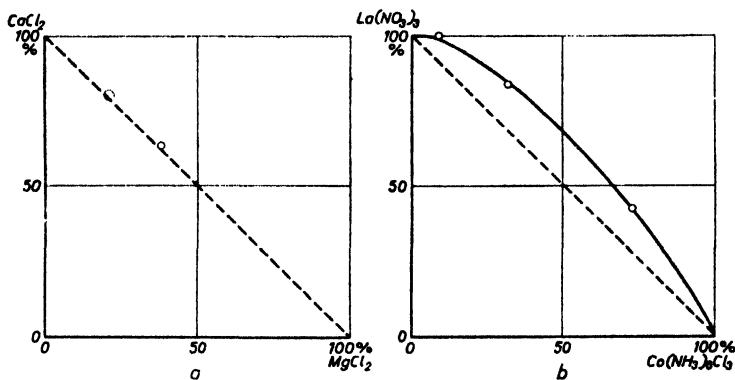
Fig. 33. Reversal of charge of alcohol insoluble soya bean phosphatide, with mixtures of $CaCl_2$+ $MgCl_2$ or $La(NO_3)_3$ + $Co(NH_3)_6Cl_3$.

[1] H. G. BUNGENBERG DE JONG, J. VAN DER MEER and L. G. M. BAAS BECKING, *Kolloid-Beihefte*, 42 (1935) 384.
[2] H. G. BUNGENBERG DE JONG and H. L. BOOIJ, *Protoplasma*, 24 (1935) 319.

A much stronger antagonism occurs in both combinations $La(NO_3)_3 + NaNO_3$, and $Co(NH_3)_6Cl_3 + NaCl$.

Now we have seen in § 2 c (p. 280), that the reversal of charge concentrations in phosphatides increase from left to right as follows:

$La^{...}$___ $Co(NH_3)_6^{...}$_____ $Ca^{..}$___ $Mg^{..}$_____ _____ $Li^{.}$___ $Na^{.}$

In this sequence we have expressed the relatively smaller differences in con-centrations by putting symbols close together.

The above given facts lead then to the idea, that the magnitude of the deviations from additivity depends on the relative positions of the cations in the reversal of charge spectrum.

If the chosen cations stand far apart (La—Na and $Co(NH_3)_6$—Na) a large antagonism, if they stand close together (La—$Co(NH_3)_6$ or Ca—Mg or Li—Na) near additivity or small deviations from that must be expected.

It has been shown convenient to use as a measure for the difference in reversal of charge concentrations the expression $\log Q = \log \frac{C_I}{C_{II}}$, in which C_I refers to the salt with the higher (salt I, e.g., NaCl or $NaNO_3$), C_{II} to the salt with the lower reversal of charge concentration (salt II, e.g., $CaCl_2$).

In Fig. 34 the results with a soya bean phosphatide fraction (the same kind as in § 5 a) are given for a number of salt combinations, which comprise the already enumerated ones in § 5 a and further the combinations hexol nitrate + $NaNO_3$; $UO_2(NO_3)_2 + NaNO_3$; $Th(NO_3)_4 + NaNO_3$ and $La(NO_3)_3 + NaNO_3$ [1].

Fig. 34. Relation between the maximum deviation from additive behaviour in the reversal of charge in salt mixtures (chlorides or nitrates of the stated ion + NaCl or $NaNO_3$) and the quotient of the reversal of charge concentrations (alcohol soluble soya bean phosphatide).
C_{II} = (true) reversal of charge con-centration of the given cation in the absence of Na salt.
C'_{II} = the same, however, in the presence of the Na salt to an amount of 42% of the reversal of charge concentration of this Na salt, if this only is present.
Q = the reversal of charge concentr-ation for the Na salt, divided by the (true) reversal of charge concentr-ation of the given salt.
If the salts behave additively C'_{II} must amount to 58% of C_{II}. The

experimentally determined points must then lie on the lowest dotted horizontal line drawn at an ordinate value of $\log C'_{II}/C_{II} = \log 0.58 = 0.76 — 1$. The combination LiCl + NaCl practically satisfies this with this phosphatide. With the other combinations antagonism occurs, i.e., the deviations from the additive behaviour are so great that the points even lie higher than the highest dotted horizontal line drawn at an ordinate value $\log C'_{II}/C_{II} = 0$, i.e. $C'_{II}/C_{II} = 1.00$ or 100%. The large complex cations (which on account of their size can no longer exert a polarising action) occupy a position apart.
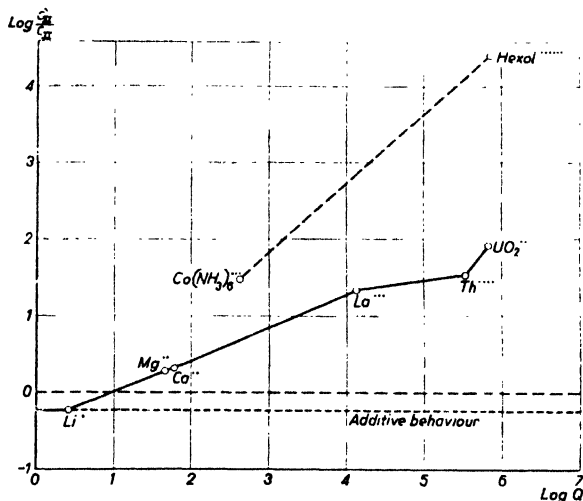
[1] H. G. BUNGENBERG DE JONG, H. L. BOOIJ and J. G. WAKKIE, Kolloid-Beihefte, 44 (1936) 254

As abscissae the above values for log Q are used and as ordinates the magnitude of the deviations from addititvity expressed by log $\frac{C'_{II}}{C_{II}}$, in which C'_{II} is the reversal of charge concentration of salt II in the presence of a concentration of NaCl (or NaNO$_3$) amounting to 42% of the reversal of charge concentration of NaCl (or NaNO$_3$). This choice corresponds to nearly the position of the maxima in the antagonism curves.

Fig. 34 reveals that the expected correlation is indeed present, but that evidently we must distinguish between the behaviour of the smaller inorganic cations and the much larger complexions Co(NH$_3$)$_6\cdots$ and the hexavalent hexol cation, the latter two showing the antagonism phenomenon on a much larger scale than the former.

This separate position of the complex cations does not surprise us so very much, as we have already seen in § 2i (p. 291) that they also stand as a class apart in the cation-spectra, representing reversal of charge behaviour.

It must be mentioned, that in constructing the above figure, the expression log $\frac{C'_{II}}{C_{II}}$ was calculated from the true reversal of charge concentrations. As we have discussed in § 2a (p. 276), the gross reversal of charge concentrations are practically true reversal of charge concentrations, if the former are sufficiently large. If they are smaller a correction for the amount of cations fixed to the colloid must be applied. This correction can be calculated from determinations of the reversal of charge with hexol nitrate at three sol concentrations, which latter determination gives also the true reversal of charge concentration of hexol nitrate, here amounting to 7.4 · 10^{-6}N. The very large antagonism in the combination hexol nitrate + NaNO$_3$ is revealed by the fact that the maximum of the antagonism curve lies here somewhat higher than 0.1 N hexol nitrate, the true reversal of charge concentration here being increased about 25 000 times (log $\frac{C'_{II}}{C_{II}} = 4.4$).

Returning to Figure 34, we see that the curve through the smaller inorganic cations intersects the level log $\frac{C'_{II}}{C_{II}} = 0$ at an abscissa value log $Q = 1$.

This means, that in salt combinations with NaCl (or NaNO$_3$) antagonism proper (p. 312) is only to be expected if the individual reversal of charge concentrations of the salts in the salt combination differ more than 10 times.

In the combination LiCl + NaCl (with $Q = 2.67$) of Fig. 32 no antagonism occurs. Now as in different phosphatide preparations the reversal of charge concentrations for the same salts may differ and therefore Q may differ, it will no longer cause surprise, that the deviations from additivity are also found to be considerably different[1]. Thus Fig. 35 gives 6 combination curves, the upper four representing combinations CaCl$_2$ + NaCl, the lower two LiCl + NaCl. On each curve the value of Q is written. We see that in the combination LiCl + NaCl also antagonism proper can occur (second curve from below), it is however here feeble, the maximum lying at only 115%. It is in accord with the above stated absolute value of Q, this being here 14, that is a value slightly above $Q = 10$.

We see further, that the antagonism becomes more and more pronounced as the value of Q increases. Combining these facts with the rôle of the reciprocal hexol number (or apparent equivalent weight) in determining the spread of cations in the

[1] H. G. Bungenberg de Jong and G. G. P. Saubert, *Protoplasma*, 26 (1936) 282.

reversal of charge spectrum, (for instance of Ca and Na, see p. 295 Fig. 20) and further with the correlation between reciprocal hexol number and purity (see p. 295 § 2 1), we may conclude that the purer the phosphatide preparation (the more it is free from phosphatidic acid, see p. 188 and 274) the more pronounced will be the anatogism $CaCl_2$ — NaCl.

Fig. 35. Reversal of charge of a number of phosphatides with mixtures of LiCl + NaCl or $CaCl_2$+ NaCl. The two lowest curves (Q = 2.7 and 14) relate to combinations LiCl + NaCl, the remainder to combinations $CaCl_2$ + NaCl.
Ordinates: concentrations of LiCl or $CaCl_2$ in the salt mixture expressed in % of the reversal of charge concentration of these salts in the absence of NaCl.
Abscissae: concentration of NaCl in the salt mixture expressed in % of the reversal of charge concentration of NaCl in the absence of LiCl or $CaCl_2$.
The deviations from the additive behaviour increase on increase of the value of Q (= quotient of the reversal of charge concentrations NaCl/LiCl or NaCl/$CaCl_2$).
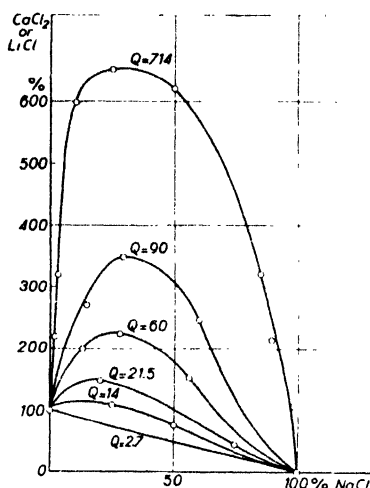
The phosphatides are colloids in which antagonism can be readily observed in the combination $CaCl_2$ + NaCl. That this is not so in other colloids of acidic
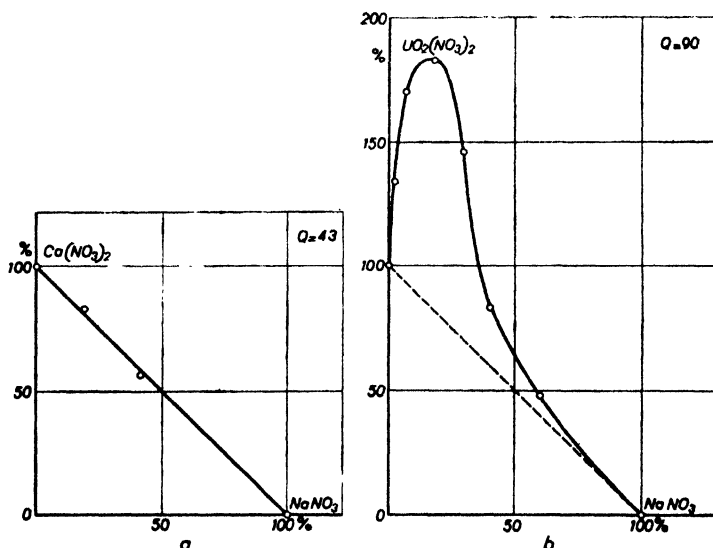
Fig. 36. Reversal of charge of Na arabinate (film adsorbed on SiO_2 particles) in mixtures of $CaCl_2$ + NaCl (a) or $UO_2(NO_3)_2$ + $NaNO_3$ (b).
The combination $CaCl_2$ + NaCl with Q = 4.27 behaves additively. In contrast with this, antagonism occurs in the combination $UO_2(NO_3)_2$ + $NaNO_3$ with Q = 90.

nature, is in accordance with the great spread of the cations in the reversal of charge spectra of phosphatides (see Fig. 10—15 and Fig. 20 on p. 281-285 and 295) and the relatively small spread in these spectra of other colloids.

Thus in gum arabic [1] the combination $Ca(NO_3)_2 + NaNO_3$ gave only additive behaviour, in accordance with the value of Q, this being $\frac{3.8}{0.89} = 4.27$ (see p. 315, Fig. 36a).

We must choose in the reversal of charge spectrum (p. 284, Fig. 14) a divalent ion which lies much further to the left if we want to obtain a distinct antagonism with gum arabic in a salt combination of the type $(2 — 1) + (1 — 1)$.

Thus indeed such an antagonism was observed in the combination $UO_2(NO_3)_2 + NaNO_3$, with an Q value of $\frac{3.8}{0.042} = 90$. (See Fig. 36b).

c. The rôle of anions in the so-called cation antagonism[2]

The above results, in which we combined salts with the same anion, seem to indicate, that if two cations are simultaneously present the cation having the larger reversal of charge concentration has an antagonistic effect on the cation with the smaller reversal of charge concentration. It seems therefore that the phenomena described could be characterised as a *"Cation Antagonism"*.

It will be shown however that this denomination is not at all an adequate one as in the antagonism the accompanying anion plays the main rôle.

We have seen in § 1 b, that reversal of charge determinations at a few sol concentrations allow to obtain two kinds of information viz. the true reversal of charge concentration and the amount of cations fixed to the ionised groups of the colloid present. Experimentally this is only possible with cations having very low true reversal of charge concentrations, the hexol cation very generally fulfilling the latter condition. Hexol nitrate was therefore very appropriate to gain information on the "charge density" of colloids (reciprocal hexol numbers, see § 1 b, p. 262).

Again for the problem interesting us here — antagonism in reversal of charge phenomena — hexol nitrate permits one to gain some insight in the underlying mechanism. Fig. 37 shows the gross reversal of charge concentrations of hexol nitrate at three sol concentrations of a soya bean phosphatide preparation without added salt (lower curve) and in the presence of some small constant concentrations of $NaNO_3$ (middle curve $= 5 \cdot 10^{-3}N$ and upper curve $13 \cdot 10^{-3}N$).

The figure reveals that the influence of added $NaNO_3$ consists in a displacement of the blank straight line towards higher concentrations, whereby its inclination is not much altered.

This permits us to conclude, that the main influence of the added $NaNO_3$ consists in a *strong increase of the true reversal of charge concentration* (the distance from intersection points of the straight lines with the ordinate axis till the origin of the coordinate system), whereas the amount of hexol ions fixed per gram of colloid (indicated by the inclination of the straight line) is not altered very much.

So in the presence of 13 m. eq. p. 1 $NaNO_3$ the true reversal of charge concentration is increased 7.9 fold, whereas the fixed amount of hexol ions (calculated from the increase of inclination) is only increased 1.37 fold.

[1] H. G. Bungenberg de Jong and G. G. P. Saubert, *Protoplasma*, 26 (1936) 282.
[2] H. G. Bungenberg de Jong, H. L. Booij and J. G. Wakkie, *Kolloid-Beihefte*, 44 (1936) 254.

Fig. 38 gives similar results for the influence of a very small constant Na_2SO_4 concentration ($1 \cdot 10^{-4}$ N), in which the antagonistic effect shows itself still more close to the ideal pure form, in which latter only the true reversal of charge concentration is strongly increased, whereas the fixed amount of hexol ions is equal to that in the blank). The reversal of charge concentration is here increased 17.5 fold, whereas the fixed amount of hexol ions is only increased 1.11 fold).
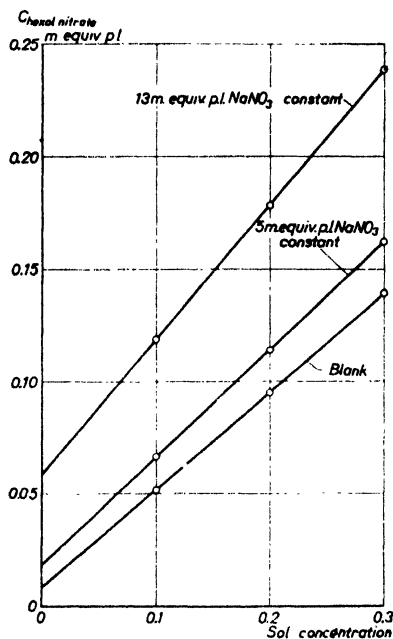
Fig. 37. Reversal of charge of alcohol soluble soya bean phosphatide with hexol nitrate as a function of the sol concentration and the influence of the presence of 5 and 13 m. eq. per l of $NaNO_3$ on it (see text).
Ordinates: (gross) concentration of hexol nitrate in m. eq. p. l.
Abscissae: sol concentration, see Fig. 4, p. 265.
Addition of $NaNO_3$ displaces the curve upwards, during which the slope changes but little (measure for the binding of hexol ions to the phosphatide); nevertheless the true reversal of charge concentration (segment which the straight line cuts off from the ordinate axis) increases considerably.
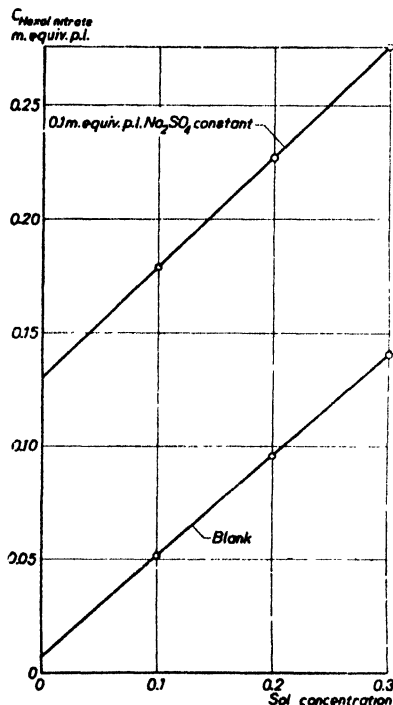
Fig. 38. Reversal of charge of alcohol soluble soya bean phosphatide with hexol nitrate as a function of the sol concentration and the influence of the presence of 0.1 m. eq. per l of Na_2SO_4 on it.
Ordinates and abscissae as in Fig. 37.
The divalent sulphate ion has a much greater influence than the monovalent NO_3 ion. The slope of the displaced curve is nearly unchanged (binding of hexol ion to phosphatide remains practically constant) but the true reversal of charge (segment which the straight line cuts off from the ordinate axis) increases very considerably.

It is further of great interest to compare the antagonistic action of Na_2SO_4 and $NaNO_3$.

We can see from Fig. 37 that 0.1 m. eq. p. l $NaNO_3$ would hardly displace the blanc curve. Replacing however the monovalent NO_3 ion by the divalent SO_4 ion

a large displacement is found. The antagonistic effect of 0.1 m. eq. p. 1 Na_2SO_4 is even larger than that of 13 m. eq. p l $NaNO_3$:

This comparison shows clearly that the denomination *"cation-antagonism"* cannot be an adequate one, the accompanying anions of the added salt playing at least an important rôle.

The equal (or slightly increasing) slope of the straight line in displacing it towards higher concentrations shows further that no sodium ions are fixed simultaneously with hexol ions, for in the latter case the slope of the straight line in displacing upwards in the figures would decrease instead.

Thus at the small $NaNO_3$ concentrations, corresponding to points on the beginning of the left ascending branch of the antagonism curve, the fixation of hexol ions is only made more difficult. This expresses itself in a considerable increase of the true reversal of charge concentrations of hexol ions needed to fix the same amount of hexol ions as in the blank.

We must therefore not seek the mechanism of this in the fixation spots of the colloid but in the surrounding medium.

Thus we are led to the idea, that the added $NaNO_3$ acts by diminishing the activity coefficient of the hexol ions in the medium. It will be clear then that the $NaNO_3$ does so mainly by its nitrate anions.

In accordance with this idea is the much stronger action of Na_2SO_4, this salt having divalent anions.

We therefore come to the conclusion, *that the ascending branch of the antagonism curve is caused practically alone by the anion of the added salt, this anion decreasing the activity coefficient of the hexol cation*[1].

d. The general form of antagonism curves and the rôle of Q in determining the intensity of deviations from additivity[2]

In choosing still larger $NaNO_3$ concentrations, the indicated process will at first continue, the true reversal of charge concentration of hexol nitrate being increased more and more.

However finally a second process becomes gradually more and more important. It will stop at last the further increase of the true reversal of charge concentration of the hexol nitrate, and will diminish the latter at very high $NaNO_3$ concentrations till it becomes zero, giving thus the antagonism curves its characteristic shape. This second process is the fixation of sodium cations on the ionised groups of the colloid, which as we saw above could be neglected at such low concentrations (5 and 13 m. eq. p. l) of $NaNO_3$, which are only a very small fraction of the reversal of charge concentration of $NaNO_3$ itself (4800 m. eq. p. l). But at higher fractions of the latter, e.g. certainly from 5 to 10 % on, the fixation of sodium next to hexol ions will no longer be negligible. Most antagonism curves show their maximum at about a fraction of 30—50 % of the reversal of charge concentration of NaCl or $NaNO_3$, thus indicating that here already the lowering effect of simultaneously fixed sodium ions compensates the increasing effect of the NO_3 or Cl anions on the true reversal of charge concentration of the hexol cation.

---

[1] A completely analogous explanation of the antagonism in the flocculation of hydrophobic colloids has been given by H. R. KRUYT, *Proc. Koninkl. Nederland. Akad. Wetenschap., Amsterdam*, 38 (1935) 464 and C. F. VESTER, *Thesis*, Utrecht 1935.

[2] H. G. BUNGENBERG DE JONG, H. L. BOOIJ and J. G. WAKKIE, *Kolloid-Beihefte* 44, (1936) 254.

At still higher fraction the lowering effect of the fixed sodium ions preponderates, the antagonism curve descends and ends at the reversal of charge concentration of $NaNO_3$ or $NaCl$, where only sodium ions are fixed.

It will be clear that the explanation given for the form of the antagonism curve hexol nitrate + $NaNO_3$ applies to other salt combinations as well, these having the general curve form in common.

It is also possible to explain why the quotient Q plays such an important rôle in determining the intensity of deviations from additivity (see p. 312, § 5b). Designating once more that salt of a salt combination with common anions, which has the higher reversal of charge concentration by I, the other salt by II, it will be evident that if these reversal of charge concentrations C_I and C_{II} lie close together, only additivity or slight positive deviations from additivity may be expected.

Taking as an example $CaCl_2$ + $MgCl_2$ in phosphatides, (see p. 295, Fig. 20) the Cl᾽ ion concentration in this combination can only increase slightly, viz. from the value C_{II} to C_I. An appreciable further decrease of the activity coefficient of the Ca ion in this combination with $MgCl_2$ is thus not possible. Further because of C_I and C_{II} lying close together at such fractions of $MgCl_2$ in the salt mixture, at which the said slight decrease of the activity coefficient might manifest itself, the Mg cation is already fixed simultaneously with the Ca ion. From both causes an almost additive behaviour in the combination curve results (see p. 312, Fig. 33a).

In the combination hexol nitrate + $NaNO_3$ just the reverse applies because of the extreme difference in the true reversal of charge concentrations of hexol nitrate $(7.4 \cdot 10^{-6} N)$ and of $NaNO_3$ (4.8N). Here the total nitrate ion concentration increases some 600 000 times in the mixtures before the reversal of charge point with $NaNO_3$ is reached. If we arbitrarily take a fraction of 5% of the reversal of charge concentration as the limit below which the fixation of Na cations is very small then on adding $NaNO_3$ up to 0.24 N an enormous increase of the total NO_3 ion concentration still occurs (some 30 000 times), which will diminish the activity coefficient of the hexol cation. The left branch of the antagonism curve can thus mount very high, before the influence of simultaneously fixed Na ions set an end to it. Having thus explained the practical additivity of $CaCl_2$ + $MgCl_2$ and the extremely pronounced antagonism in the combination hexol nitrate + $NaNO_3$, no further discussion of the other salt combinations lying in between these extreme cases will be needed. It will be clear that the experimentally found importance of $Q = \dfrac{C_I}{C_{II}}$ for the magnitude of the positive deviations is thus self evident. The lower curve in Fig. 34 (p. 313) shows the correlation which might be expected from the above discussion. The isolated position of the voluminous complex cations may be explained by a relatively much larger decreasing influence of anions upon their activity coefficient than upon that of the small inorganic cations.

We have already discussed the influence of the anion valency in § 5c, Na_2SO_4 showing a much larger influence in increasing the true reversal of charge concentration of hexol nitrate than $NaNO_3$.

It must be expected that this valency influence will also express itself in the height of the maximum of the antagonism curve.

This indeed was shown to be the case in comparing the combinations $CaCl_2$ + $NaCl$ and $CaCl_2$ + $Na_2S_2O_3$ [1], the maximum in the latter case lying about 5 times as high as in the former.

---

[1] Here the divalent thiosulphate anion was chosen, as with sulphate precipitation of $CaSO_4$ would occur.

e. Reversal of charge with mixture of three salts[1]

On three different phosphatide preparations the reversal of charge was investigated in mixtures of three salts. ($CaCl_2$ + $MgCl_2$ + NaCl and La $(NO_3)_3$ + $Co(NH_3)_6Cl_3$ + NaCl).

The results, which might be of interest for certain biological problems, do not open further theoretical perspectives and will therefore be not discussed here.

f. Reversal of charge with mixtures of salts containing organic cations
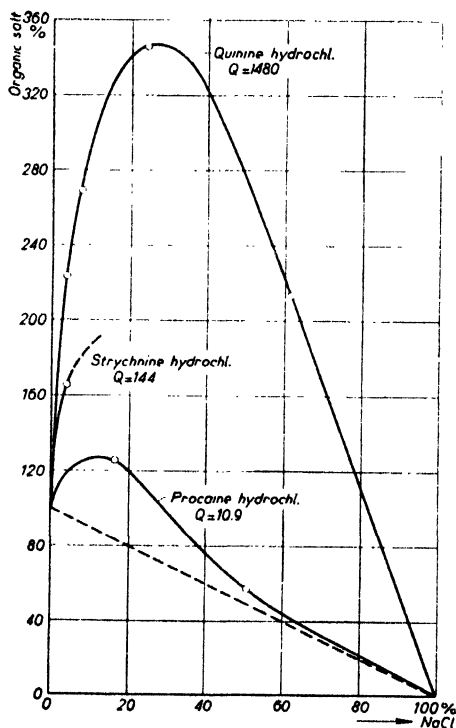
As yet we have discussed only mixtures of inorganic salts. Antagonism effects may be expected and have indeed been found to exist in combinations in which organic salts take part[2].

Fig. 39 shows the results obtained in a soya been phosphatide sol for the combinations of three alkaloid chlorides each with NaCl, here also the antagonism being more pronounced as Q increases, that is the greater the distance ($=$ log Q) between the reversal of charge points in the ion spectrum (see p. 301, Fig. 24).

In these combinations with NaCl the above found rule that antagonism proper can be expected if Q is larger than 10 (that is a distance in the ion spectrum of one logarithmic unit) seems to hold here also.

Fig. 39. Reversal of charge of alcohol soluble soya bean phosphatide with mixtures of quinine hydrochloride + NaCl; strychnine hydrochloride + NaCl and procaine hydrochloride + NaCl. The deviations from additive behaviour increase as Q rises. In this case the rule still holds that antagonism (i.e., elevation of the curve above the level of 100%) occurs when Q exceeds the numerical value of approximately 10.

But from the results of the combinations of quinine, strychnine and procaine hydrochlorides each with $CaCl_2$, a different behaviour was found, indicating that here the value of Q, necessary for obtaining antagonism proper is much higher. (See Fig. 40 a and b).

Still in these three combinations the rule holds, that the positive deviations from additivity increase with increasing Q, but with the highest value of $Q = 41,6$ (quinine hydrochloride-$CaCl_2$) still no antagonism proper is obtained.

[1] H. G. BUNGENBERG DE JONG, J. VAN DER MEER and L. G. M. BAAS BECKING, *Kolloid-Beihefte*, 42 (1935) 384; H. G. BUNGENBERG DE JONG and H. L. BOOIJ, *Protoplasma*, 24 (1935) 319.
[2] H. G. BUNGENBERG DE JONG and J. G. WAKKIE, *Biochem. Z.*, 297 (1938) 70.

This result seems to indicate that no general importance can be ascribed to $Q = 10$, the necessary value for obtaining antagonism proper depending not only on the quotient of the reversal of charge concentrations but also on the absolute height of the reversal of charge concentration of salt I. NaCl has always a relatively very high reversal of
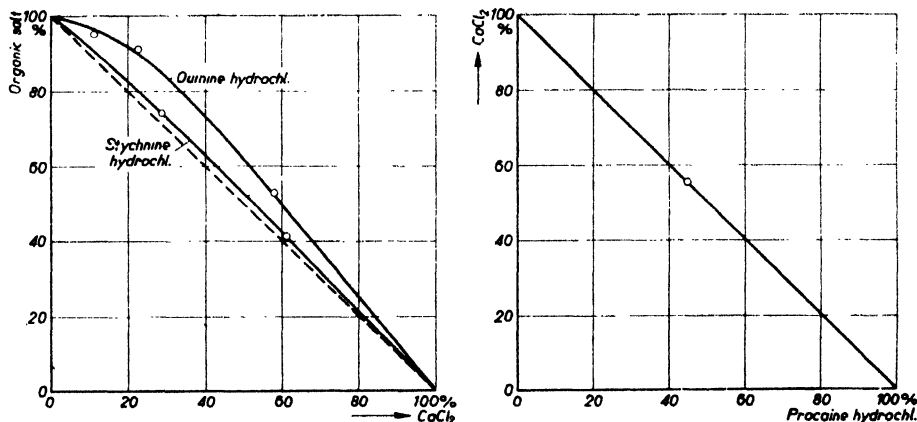
Fig. 40. Reversal of charge of alcohol soluble soya bean phosphatide with mixtures of quinine hydrochloride + CaCl₂ (Q = 41.6), strychnine hydrochloride + CaCl₂ (Q = 4.1) and procaine hydrochloride + CaCl₂ (Q = 3.25).
The deviation from the additive behaviour increases as Q rises, but in these combinations with CaCl₂ the rule, that antagonism begins when Q exceeds the value 10, no longer holds (see text).

charge concentration, $CaCl_2$ in phosphatides a much lower one. In combinations with $CaCl_2$ therefore the Cl ion concentration never can attain such high absolute values as in combinations with NaCl. This may explain why for obtaining antagonism proper the necessary value of Q is much higher in the combinations with $CaCl_2$ than in those with NaCl. It is to be expected that this will not only apply for combinations of alkaloid salts with $CaCl_2$ or NaCl but also for those of inorganic salts with $CaCl_2$ or NaCl.

§ 6. REVERSAL OF CHARGE PHENOMENA IN MIXTURES OF COLLOIDS

a. Reversal of charge at constant pH on varying the mixing ratio of oppositely charged colloids

Interaction between oppositely charged sols is a general phenomenon in colloid science, occurring both in "hydrophobic" and "hydrophylic" colloids. This interaction usually manifests itself in flocculation, the latter being a maximum at or near to the mixing ratio of the sols corresponding to the reversal of charge point of the floccules.

BUNGENBERG DE JONG and coworkers found this also to hold in those cases of "hydrophilic" colloids, in which coacervation occurs ("complex-coacervation", see p. 338, Chapter X, § 2). An example is given in Fig. 41, giving the electrophoretic

velocities of the small coacervate drops at different mixing ratios of 0.15% gelatin and gum arabic sols at pH 3.50 [1].

The electrophoretic velocities of each of the sols separately and of the clear mixture at a ratio of 83.3% gum arabic sol were measured with suspended carborundum particles, these latter serving in the same way as SiO_2 particles as an "electrophoretic indicator". The figure shows clearly the interaction of the positive and negative colloid, the electrophoretic velocity being intermediate between those of the gelatin and of the gum arabic, which leads necessarily to the occurrence of a reversal of charge point. Especially we lay stress on the fact, that the electrical interaction is not limited to the actual occurrence of coacervation, for also in those ranges of ratio in which no coacervation occurs (dotted parts of the curve) the electrophoretic velocity decreases on addition of the second colloid.
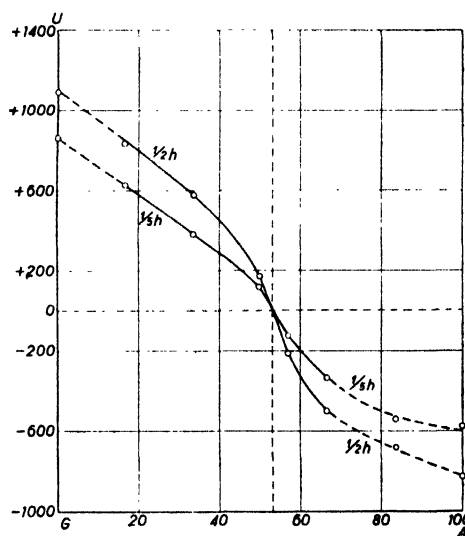
Fig. 41. Electrophoretic velocity of coacervate drops or carborundum particles in mixtures of isohydric 0.05% gelatin (G) and gum arabic sols (A).
Ordinates: U electrophoretic velocity at $^1/_6$ or $^1/_2$ cell height expressed in arbitrary units.
Abscissae: Mixing proportion of the isohydric sols, expressed in % of gum arabic sol in the total mixture.
The dotted portions of the curves (clear mixtures measured with SiC particles) join on to the continuously drawn portion (measured on small separated coacervate drops), which points to the formation of a complete film containing both colloids on the surface of the SiC particles. The intersections of the two curves (measurements at $^1/_6$ and $^1/_2$ cell height respectively) with the level $U = 0$ coincide, which shows that the glass wall also is covered with a complete film, the composition of which is the same as that present on the surface of the coacervate drops (interaction product of G and A).

b. Reversal of charge at constant mixing ratio on varying pH

Fig. 42 shows the electrophoretic velocity as a function of pH of a number of mixures of 0.05% sols of gelatin and gum arabic [2]. The two dotted vertical lines give the pH at which each of the colloids used has its reversal of charge point, viz., gelatin 4.82 and gum arabic 1.73 [3].

---

[1] H. G. BUNGENBERG DE JONG and W. A. L. DEKKER, Kolloid-Beihefte, 43 (1935), 143 cmf. p. 132.

[2] H. G. BUNGENBERG DE JONG and W. A. L. DEKKER, loc. cit., p. 185 and 190.

[3] J. G. WAKKIE, Thesis, Leiden 1936, p. 36. One would not a priori expect a reversal of charge point with pH for gum arabic, this being the salt of a high molecular carboxylic acid. It seems from the point of view of dissociation more probable, that the electrophoretic velocity would decrease asymptotically to zero. Indeed the curve has that form, but the level lies at a very small positive value. Perhaps this slight positive charge is due to an impurity containing nitrogen (0.33% N) with could not be removed by purification.

The figure shows that at each mixing ratio a definite pH value exists for the reversal of charge point, which lies in between that of gelatin and gum arabic. Putting these reversal of charge points in a diagram, as a function of the mixing ratio and pH, we obtain Fig 43.

If we assume, that negative gum arabic and positive gelatin, when present together do not in reality occur free, but combine by virtue of their opposite charges,
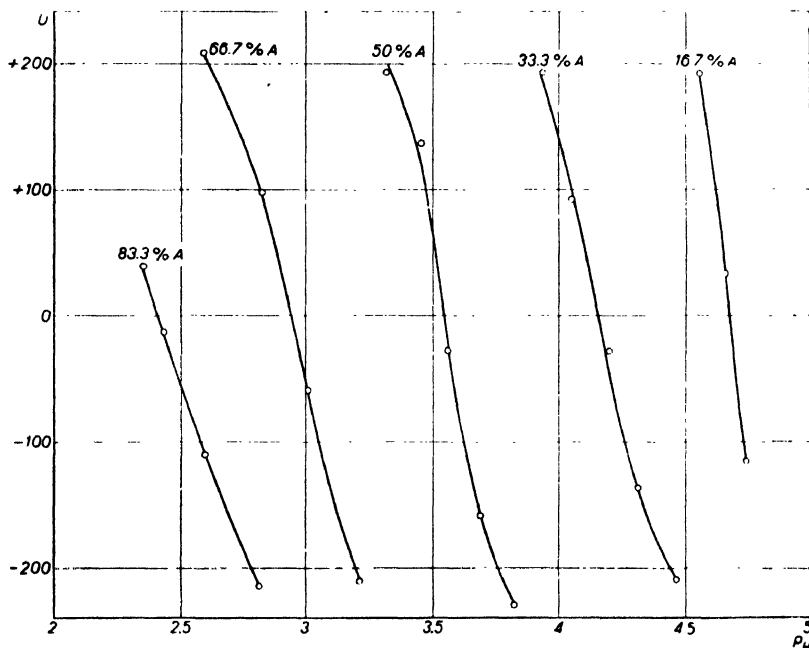
Fig. 42. Reversal of charge in mixtures of 0.05% gelatin and gum arabic sols on variation of the pH. Each curve relates to a constant mixing proportion of the sols. The mixing proportion is expressed in % of the gum arabic sol (A) in the mixture of the two sols.
Ordinates: U = electrophoretic velocity of the coacervate drops produced, expressed in arbitrarily chosen units.
Abscissae: pH.

the reversal of charge point must be found at that mixing ratio, where gelatin and gum arabic are present in equivalent quantities. The mixing ratio must thus depend on the equivalent weights of each of the colloids. It is perhaps better to use here the term apparent equivalent weight (see p. 273, § 1f), for in the pH range here con sidered we have neither to do with the minimum equivalent weight of positively charged gelatin (which is only reached below pH 3) nor with that of gum arabic (which is reached above pH 5).

Both apparent equivalent weights are functions of pH and to get an impression of the relative change of these equivalent weights with pH, we can use the electrophoresis pH curves of both colloids. See Fig. 44.

At the *I. E. P.* of the gelatin used (pH 4.82), the electrophoretic velocity being zero, the apparent equivalent weight of gelatin is infinite. That of gum arabic has a finite value, not being very different from its value at pH 6.
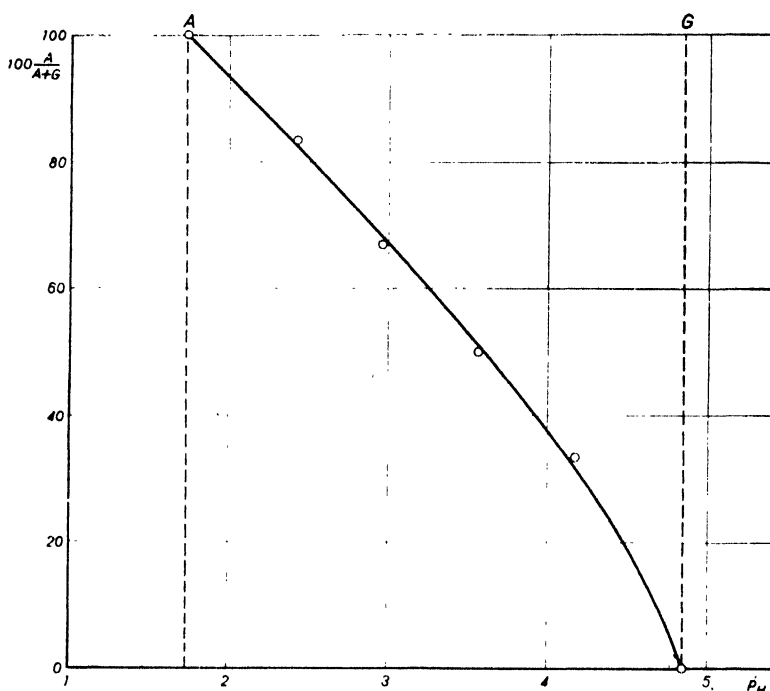
Fig. 43. Position of the reversal of charge in a mixing proportion — pH diagram (according to the data of Fig. 42).
Ordinates: mixing proportion in °/₀ A (as in Fig. 42) of the 0.05% gelatin (G) and gum arabic (A) sols.
Abscissae: pH.
The two vertical dotted lines give the pH values at which the colloid preparations used have the electrophoretic velocity = 0. Between these vertical lines the two colloids therefore carry charges of opposite sign (gelatin positive and gum arabic negative). The reversal of charge points lie on a curve, which stretches from pH 1.73 at 100% A to pH 4.82 at 100% G.

Now considering a slightly lower pH than the *I.E.P.* of gelatin, this colloid acquires a small free positive charge, so that its equivalent weight has become finite, though still very much larger than that of gum arabic (the equivalent weight of the latter has hardly changed by the slight pH decrease, see the flat course of the electrophoresis curve near pH 4.8). Thus to combine equivalent amounts of gelatin and gum arabic we need a large quantity of gelatin and a small one of gum arabic, *i. e.*, the mixing ratio of the reversal of charge expressed in % A will be small.

If the pH is decreased further the changes in apparent equivalent weights proceed further in the directions indicated (that of gelatin decreases, that of gum

arabic increases), so that the mixing ratio for reversal of charge continually shifts towards higher % A. Having thus explained why the curve in Fig. 43 must start from the *I. E. P.* of gelatin at a mixing ratio of 0% A, we need not repeat in detail an analogous reasoning why the curve must end at the reversal of charge point of gum arabic at 100% A (the apparent equivalent weight of the latter here being infinite, that of gelatin however, finite).

What has been said above on the reversal of charge in a mixture of gelatin and gum arabic, must of course apply for a positive and a negative colloid in general. Thus also for the case of a mixture of two proteins, in the pH range between the two I. E. P.
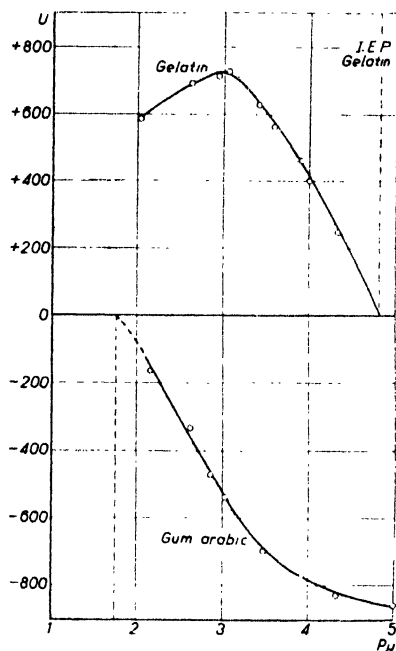
Fig. 44.   Electrophoretic velocities of 0.67% gelatin or gum arabic sols as a function of the pH.
Ordinates: U = electrophoretic velocities of the carborundum particles suspended in the sols, expressed in arbitrarily chosen units.
Abscissae: pH.

The particular form of curves similar to that of Fig. 43 must in principle be calculable from data on the apparent equivalent weight (see p. 273) of both colloids involved.

It can be foreseen that such curves will not necessarily always show exactly the same shape, for this will depend on the variation of the apparent equivalent weights with pH.

If for instance two colloids are used, which show a practically constant equivalent weight in a certain pH range, then in curves analogous to Fig. 43, a horizontal tract must occur, that is the combining ratio of the two colloids must in a certain pH range be independent of pH. The electrophoresis — pH curve of clupein, a basic protein, shows in a certain pH range an electrophoretic velocity independent of pH (see Fig. 45) [1]. Thus it may be expected that in certain colloid mixtures, in which clupein is taken as the positive component, the above mentioned form of curve may occur.
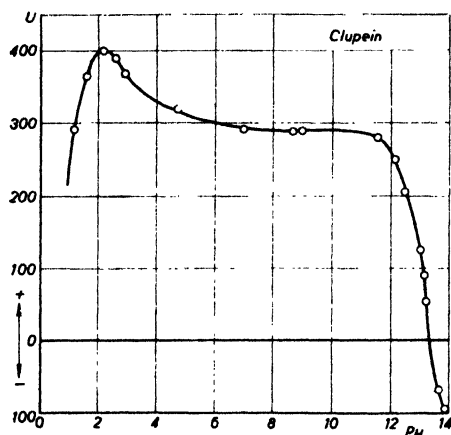
Fig. 45.   Electrophoretic velocity — pH curve of clupein.
Abscissae: pH.
Ordinates: electrophoretic velocity of SiO_2 particles suspended in the 0.06% sol, expressed in arbitrarily chosen units. Since it follows from experimental observed coating curves at both pH = 2 and pH = 11 that the clupein film on SiO_2 particles is already complete from 0.006% clupein onwards, the chosen sol concentration signifies an at least 10 fold margin of safety.

[1] L. TEUNISSEN-VAN ZIJP, *Thesis*, Leiden 1938, p. 24.

A second possibility for such curve forms is of course if in a certain pH range the equivalent weights of both colloids are changed (both decreased or both increased) in such a way that their ratio is not altered.

A simultaneous increase in apparent equivalent weight of both colloids is already present in Fig. 44 in the pH range from pH 3 — 1.73 but it does not manifest itself in Fig. 43 as a horizontal part in the curve, because the percentage change in the appearent equivalent weight of gum arabic (reaching infinity at pH 1.73) is very much larger than that of gelatin. A horizontal part in the curve of Fig. 43 could only be expected if the percentage changes were exactly equal.

c.   The reversal of charge in a mixture of oppositely charged colloids as a function of the colloid concentrations

In the preceding subsection the supposition was made that positive and negative colloids at the reversal of charge point combine in equivalent proportions, practically not leaving either of the colloids free in equilibrium with the colloid-colloid combination.

We will first consider this supposition in more detail and then describe experiments which are in favour of it.

The above mentioned supposition has its logical foundation in the results obtained with organic ions in § 4 (p. 300). We saw there that the reversal of charge of colloids with organic ions shows in the first place a certain simplification compared to the reversal of charge with small inorganic ions, in so far as polarization phenomena no longer play a rôle in the fixation of ions on the ionised groups of colloids.

The second important result was that both in negative and in positive colloids the true reversal of charge concentrations decrease as the organic cation or anion increases in size.

Now considering that the ionised kinetic units of colloids are indeed very large organic ions, it seems logical to assume, that the behaviour of an oppositely charged colloid towards a given colloid will not be principally different from that of an oppositely charged large organic ion. It must be expected first, that every positively charged colloid can bring about reversal of charge of every negative colloid, specificities originating from different polarisabilities being absent.

As both components of a mixture of oppositely charged colloids can be regarded as a very large organic ion acting every time on an oppositely charged colloid, the true reversal of charge concentration of both colloids must further be expected to be very small. That means that the colloids must combine at the reversal of charge point in equivalent proportion, leaving practically neither of the colloids free in equilibrium with that colloid-colloid combination. Fig. 46 shows data calculated from experiments[1] on the reversal of charge at constant pH by varying the mixing-ratio of a number of equally concentrated gelatin and gum arabic sols. The experimental points give the composition in the sol mixture (content of gelatin or gum arabic in %) at the reversal of charge. The two points at small colloid concentrations are taken from another series of measurements. Inspecting the figure we see that the experimental points lie on or near to a straight line through the origin.

The accuracy of the experiments was however not ideal. The irregular position of the points at higher colloid concentration are possibly influenced by the need of

[1] H. G. Bungenberg de Jong and W. A. L. Dekker, *Kolloid-Beihefte* 43 (1936) 213, Cmf. p. 242-243.

measuring here the electrophoretic velocity in systems containing large coacervate drops. Inherent difficulties in working with gelatin systems (electrophoresis measurements at 40°, changes of gelatin with time) interfere also with the accuracy of these experiments. Still the results of Fig. 46 point in the direction that positive gelatin and negative gum arabic combine in practically equivalent proportions, and that the equilibrium concentration of either of the colloids at the reversal of charge point is negligibly small.

Fig. 46. Reversal of charge of gum arabic with gelatin at pH 3.5 as a function of the gum arabic concentration.
Ordinates: gelatin "concentration" in the total system in %.
Abscissae: gum arabic "concentration" in the total system in %.
At higher colloid concentrations technical difficulties occur in the measurement of the electrophoretic velocity, to which the irregular position of the points is attributed The points are scattered around a straight line through the origin (see text).
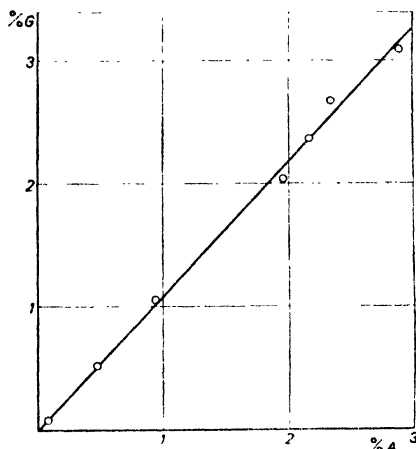
With a more accurate technique and using a colloid combination in which the above difficulties do not exist, L. TEUNISSEN-VAN ZIJP [1] showed this indeed to be the case for clupein sulphate and Na-arabinate sols.

Fig. 47 shows the results. The three experimental points lie on a straight line intersecting the ordinate axis at an extremely small concentration ($8 \cdot 10^{-5}$ %). As the equivalent weight of Na arabinate is known from analysis (1202, see p. 262, § 1 b) this equilibrium concentration (if at all real) would only be $7 \cdot 10^{-7}$ N.
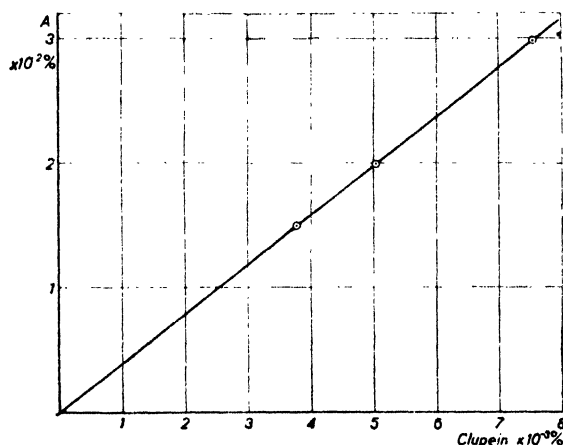
Fig. 47. Reversal of charge of clupein with Na arabinate at a pH of about 7-8, as a function of the clupein concentration.
Ordinates: arabinate "concentration" in the total system in 10^{-2} %.
Abscissae: clupein sulphate "concentration" in the total system in 10^{-3} %.
The straight line passes through the origin of the graph, i.e. at the reversal of charge point all arabinate anions are bound to the clupein cations. From the slope of the straight line and the equivalent weight of Na arabinate one can calculate a "reciprocal arabinate number" of 304 for clupein (see text)

---

[1] L. TEUNISSEN-VAN ZIJP, Thesis, Leiden 1938, cf. p. 44.

From the slope of the straight line we can calculate with the aid of the said equivalent weight of Na arabinate the "reciprocal arabinate number" of clupein (analogous to calculations of reciprocal hexol numbers of negative colloids, see § 1 *b*). 304 is then found for this number. To check this value for the electrophoretically found charge density, it can be compared with the charge densities determined with germanin (Bayer 205), the sodium salt of a hexavalent sulphonic acid. This salt proved very useful for determining reciprocal ion numbers of positively charged colloids, the equilibrium concentrations being sufficiently small (smaller than 10^{-6} N).

The reciprocal germanin numbers of clupein were determined at two pH values, at pH 2.1 and pH 11, and were found to be 220 and 261 resp. These pH values were originally chosen as a test of the supposition already made use of in § 6b, that the electrophoretic velocity of very dilute sols is an inverse measure of the equivalent weight. Inspecting Fig. 45 (p. 325), we must thus expect that the equivalent weight at pH 2.1 is smaller than at pH 11. The results confirm this assumption.

If at pH 7—8 the reciprocal germanin number had been determined, a value very near to 261 would have been obtained (because of the nearly horizontal level of the U-pH curve in Fig. 45 between pH 7—11).

Taking into account that the reciprocal germanin number is 10—15% smaller than the equivalent weight (see § 1 *c*, p. 269); the (apparent) equivalent weight of clupein at this pH would amount to about 290—307.

Now at that pH a reciprocal Na arabinate number of 304 was found, that is, practically the same value as the equivalent weight calculated from the reciprocal germanin-number.

From this we may conclude, that also quantitatively a charged colloid (here Na arabinate) behaves towards an oppositely charged one (here clupein) in the same way as a large organic ion (here germanin), but with the simplification that the systematic difference between reciprocal anion number and equivalent weight is much smaller (or perhaps absent) if the anion in question is a colloid anion.

d. The reversal of charge in mixtures of three colloids

The reversal of charge in a mixture of three colloids has been investigated in the system gelatin (positive) + Na arabinate (negative) + Na nucleate (negative) at a pH of approximately 3.6[1].

The results are given in a triangular diagram, see Fig. 48, which as is well known is a convenient method of plotting the composition of a mixture of three components, the sum of the latter being constant. The diagram is here used to plot the mixing ratios of the three individual sols in question. Thus the corners G, A and N represent these sols themselves. Points on the three sides of the diagram represent mixtures of two sols and points within the plane of the triangle mixtures of the three sols. As the sols were chosen nearly equally concentrated (1.3%), the mixing ratios of the sols coincide practically with the mixing ratios of the colloids themselves. No reversal of charge point was found on the side A N, this of course not being expected for a mixture of these two negative sols.

In the binary mixtures gelatin + arabinate (side GA) and gelatin-nucleate (side GN) reversal of charge points occur. The mixing ratio at which the reversal

[1] H. G. BUNGENBERG DE JONG and E. G. HOSKAM, *Proc. Koninkl. Nederland. Akad. Wetenschap.,* Amsterdam, 45 (1942) 387.

of charge point of the latter combination is situated, expressed in % N is lower (nearer to the corner G), then the mixing ratio in the combination gelatin + arabinate, expressed in % A (further away from the corner G). This is in accord with the lower equivalent weight of the nucleate. It has now been found that the reversal of charge points within the plane of the triangle lie practically on a straight line connecting the points of reversal of the two sides of the triangle GA and GN. This line thus divides the plane of the triangle into two parts, in an upper half up to the corner G in which the systems are positive, and in a lower half in which the systems are negative.
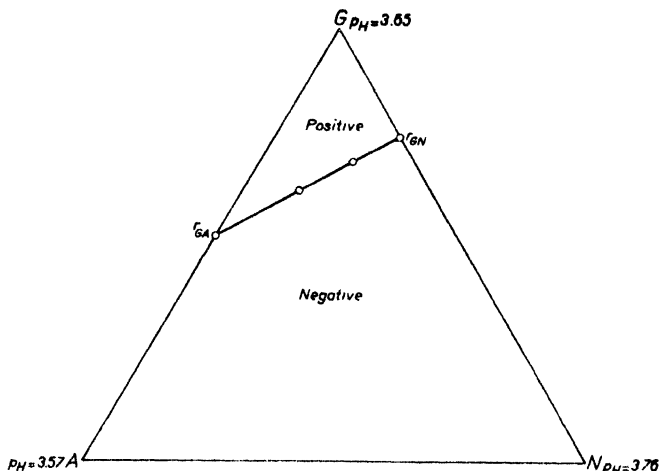
The complicated coacervation phenomena occurring within the plane of the triangle at mixing ratios in the neighbourhood of the reversal of charge line will be further discussed in Chapter X § 2t (p.378).

Fig. 48 gives a survey of the reversal of charge in mixtures of the three colloids for one pH only.

Fig. 48. Reversal of charge in mixtures of equally concentrated isohydric sols of gelatin (G), gum arabic (A) and nucleate (N). At the chosen pH G is positive, but A and N are negatively charged. The mixtures containing G + A or G + N are represented by points on the sides of the triangle GA and GN respectively. At every time one of these points is the reversal of charge point. Since the equivalent weight of N is appreciably lower than that of A, the reversal of charge point rGN on the side GN lies closer to the corner G than the reversal of charge point rGA on the side GA. Mixtures within the plane of the triangle represent mixtures which simultaneously contain G, A and N. The reversal of charge points of these mixtures lie on a straight line which joins the reversal of charge points cn the sides GA and GN.

Reversal of charge dependent both on mixing ratio of the three colloids and pH, can no longer be represented in a two dimensional diagram as was the case in mixtures of two colloids (p. 324, Fig. 43). We can then use a prism erected on the triangle GAN as base, the pH being set out along its ribs. The reversal of charge points in such a space diagram lie on a surface, which as is easily seen must intersect the rib erected on G at the I.E.P. of G.

For, this reversal of charge surface originates from the displacement of the reversal of charge line in Fig. 48 by changing the pH. Now this displacement depends on the displacement of its end-points on the sides GA and GN, these being the reversal of charge points in the binary combinations G + A and G +N. Both points will shift along the sides GA and GN away from G by lowering the pH, and in the direction of G by increasing the pH (see Fig. 43). As in the latter case for both combinations GA and GN, the reversal of charge point will reach the value 100%G at the same pH, namely at the I.E.P. of G, the reversal of charge surface must necessarily intersect the rib erected on G at the I.E.P. of G.

e. Retrograde reversal of charge phenomena

The reversal of charge phenomena in a mixture of colloids by changing pH as discussed in § 6b (p. 321) may be considered as the normal case. The colloid mixture containing oppositely charged colloids behaves as a new colloid having an "intermediate *I.E.P.*", the position of this reversal of charge point depending on the mixing ratio of the colloids.

For every such mixture the shape of the reversal of charge curve is in principle the same, that is the curve starting from negative values mounts to positive values, intersecting the reversal of charge level only *at one point*, the reversal of charge point of the given mixture (see p. 323, Fig. 42).

We may now ask if this applies also for a mixture of negative colloids if not pH is altered, but the concentration of an added salt is increased, which salt is supposed to be capable of bringing about reversal of charge of each of the colloids separately.

This problem has not been investigated systematically, only two examples having been studied. In the first [1] (soya bean phosphatide + gum arabic, and $La(NO_3)_3$ as salt) a behaviour in principle analogous to gelatin + gum arabic on changing the pH was found. The second [2] case studied — soya bean phosphatide + Na nucleate and $CaCl_2$ as salt — gave results which differ so much from "normal behaviour" that its discussion here seems quite justified. Fig. 49 gives the electrophoretic velocity as a function of the log. of the $CaCl_2$ concentration for a 0.136% soya bean phosphatide sol (A), a 0.085% sodium nucleate sol (F) and for a series of mixtures of both sols (B,C,D and E, containing respectively 9.1; 50; 90.9 and 99% of the nucleate sol).
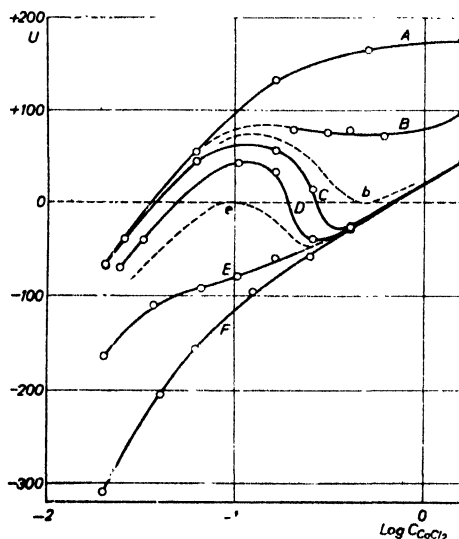
Fig. 49. Electrophoretic velocity ot ΓiO_2 particles, suspended in 0.136% soya bean phosphatide sol (A), in 0.085% Na nucleate sol (F) or in mixtures of these sols (B, C, D, E), as a function of the $CaCl_2$ concentration.
Ordinates: U = electrophoretic velocity expressed in arbitrarily chosen units.
Abscissae: Logarithm of the $CaCl_2$ concentration in equiv. per l.
B, C, D and E contain respectively 9.1, 50, 90.9 and 99% of the nucleate sol.
A, B, E and F exhibit only one reversal of charge point, C and D however three reversal of charge points, of which each time the middle one is retrograde (see text).

As a flocculation occurred neither in the phosphatide nor in the nucleinate sol nor in any of their mixtures, the electrophoretic velocity was measured on sus-

[1] H. G. BUNGENBERG DE JONG and G. G. P. SAUBERT, *Bioch. Z.*, 288 (1936) 1.
[2] H. G. BUNGENBERG DE JONG and H. I. JOUKOVSKY, *Comptes rendus des séances de la société de biologie*, 123 (1936) 511.

pended TiO_2 particles (serving as "electrophoretic indicator" in the same way as SiO_2 is used for that purpose)[1].

The curves for the sol mixtures show a very abnormal course, especially the curves C and D are remarkable in that they show *three reversal of charge points*. Taking curve C (ratio 50%) and starting from the negative sol mixtures addition of $CaCl_2$ first brings about reversal of charge at a $CaCl_2$ concentration only slightly higher than for the phosphatide sol alone. The curve mounts to a certain positive value of the charge but further addition of $CaCl_2$ decreases this positive charge and brings about once more reversal of charge, the charge attains a certain negative value and on further increasing the $CaCl_2$ concentration for the third time reversal of charge is obtained. This third reversal of charge point lies very close to that of Na nucleate alone. In the middle range of $CaCl_2$ concentrations (from log C $= 0.00 — 1$ up to $0.40 — 1$, that is from 0.1 N up to 0.25 N) the U-log C curve takes a very abnormal course, the positive charge is here changed to a negative one in increasing the $CaCl_2$ concentration. This can be designated as a "*retrograde reversal of charge*", for increasing the $CaCl_2$ concentration brings about reversal of charge from negative to positive in each of the colloids separately, which latter is the normal behaviour.

It can be further seen from Fig. 49 that there must also exist two mixing ratios of the colloids which possess one normal reversal of charge point and a point at which the charge reaches the value zero without reversal of charge to the opposite sign following.

They are the dotted curves b and e in the figure, the probable positions of which have been obtained by interpolation. If we compare Fig. 49 with Fig. 42 (p. 323) the quite different behaviour of the reversal of charge phenomena in nucleate + phosphatide mixtures with $CaCl_2$ and in arabinate + gelatin mixtures with pH is striking. This totally different behaviour is also obvious by comparing Fig. 43 (p. 324) with Fig. 50 (p. 332).

The latter figure gives the position of the reversal of charge points in Fig. 49 as a function of the mixing ratio in % nucleate and of the logarithm of the $CaCl_2$ concentration.

In order to make this Fig. 50 in all respects comparable with Fig. 43 the mixing ratio is expressed as g nucleate divided by g nucleate + g phosphatide (calculated from the mixing ratios of the sols and their concentrations), further the logarithm of the $CaCl_2$ concentration is taken in Fig. 50 increasing from right to left (in Fig. 43 the logarithm of H concentration increases from right to left, the pH being the negative logarithm of the H concentration).

The behaviour of a given colloid mixture, for instance the 50% mixture, can be read from the figure by drawing a horizontal line at that colloid-ratio. Thus the line representing the 50% mixture (dotted horizontal line) intersects the curve at a, b and c successively in increasing the $CaCl_2$ concentration. Of these three reversal of charge points a and c are normal reversal of charge points, b is the retrograde reversal of charge point.

---

[1] In 0.001% sols containing only phosphatide or only nucleate a complete film on the TiO_2 particles is formed at the respective reversal of charge points with $CaCl_2$. In all above sol mixtures B, C, D, E the concentration of each colloid — even of nucleate in B and of phosphatide in E — is larger than 0.001%.

We see from the figure that in a large range of mixing ratios of the colloids (21— 96% N) these three reversal of charge points occur, this being a quite different behaviour from that in the colloid combination gum arabic + gelatin, where at every colloid ratio only one reversal of charge point occurs.
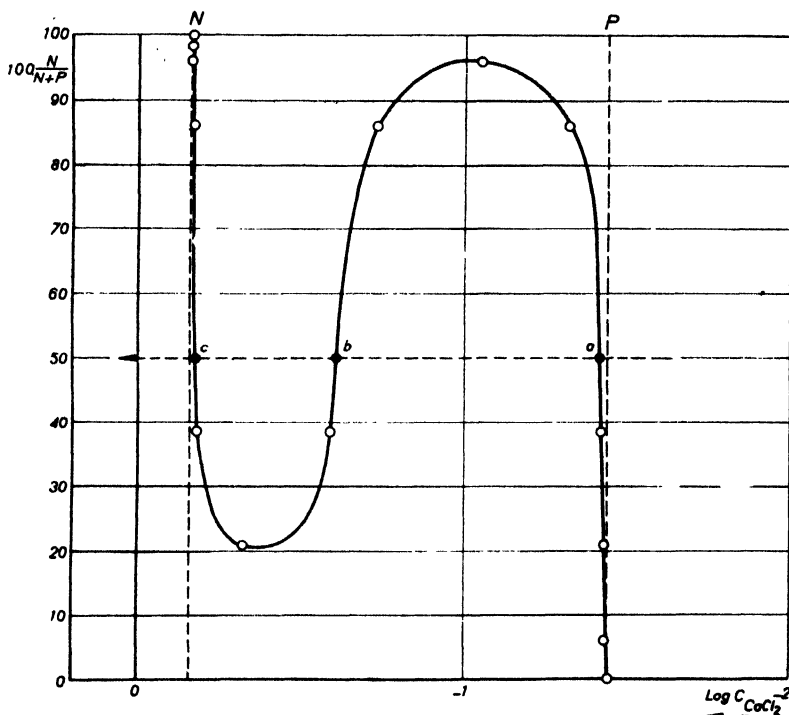
Fig. 50. Position of the reversal of charge points in a mixing proportion — log C_{CaCl_2} diagram according to date of Fig. 49.
Ordinates: colloid ratio of nucleate in the nucleate-phosphatide mixture (calculated from the mixing proportion of the sols and concentration of the sols).
Abscissae: logarithm of the $CaCl_2$ concentration; to make the figure as readily as possible comparable with Fig. 43 (p. 324), the abscissae are plotted rising from right to left (thus the $CaCl_2$ concentration falls from left to right, just as in Fig. 43 the H ion concentration falls from left to right).
Dotted vertical lines: situation of the reversal of charge points of nucleate or phosphatide with $CaCl_2$. For horizontal dotted line and intersections a, b, and c with the curve see text.

The simplest explanation for these striking differences between Fig. 43 and 50 seems to be that whereas in the combination positive gelatine + negative gum arabic colloid-colloid interaction occurs, this interaction is almost lacking in the combination positive phosphatide + negative nucleate. The 50% mixture behaves at the smaller $CaCl_2$ concentrations as if practically only the phosphatide is present, at the higher $CaCl_2$ concentrations as if only the nucleate is present,

the reversal of charge point *a* lying very close to the reversal of charge point of phosphatide itself and the reversal of charge point *c* lying very close to that of nucleate itself.

The behaviour of the 50% mixture can be explained hardly otherwise than by the absence of colloid-colloid interactions and the preferential formation of a colloid film on the TiO_2 particles of just that colloid component of the mixture of free colloids, which at the given $CaCl_2$ concentration is near to its reversal of charge point.

It seems plausible to assume that the formation of a complete colloid film is optimal just at the reversal of charge point of a colloid[1]. Evidently this plays an important rôle in the competition of the phosphatide and nucleate. Thus at *a* the colloid film on the adsorbed particles consists practically only of phosphatide, at *c* only of nucleate. The remarkable retrograde reversal of charge (point *b*) on increase of the $CaCl_2$ concentration can thus be seen as result of the gradual replacement of positive phosphatide by negative nucleate.

To explain the actual form of the curve in Fig. 50, we have to consider, that apart from the influence of the charge state of course also the influence of the colloid concentration on film formation (p. 277, § 2b) must be of importance.

For in changing the colloid ratio, we necessarily also alter the absolute concentrations of both colloids in opposite directions.

At a ratio of 86% N, the phosphatide concentration being smaller, the nucleate concentration however larger than at 50%N, the removal of the positive phosphatide film in favour of a negative nucleate film will thus be found facilitated, that is the retrograde reversal of charge point will be reached at a smaller $CaCl_2$ concentration (this point lies therefore in Fig. 50 more to the right than point b). The relative excess of nucleate already manifests itself even at smaller $CaCl_2$ concentrations. The phosphatide film contains here a much larger amount of nucleate than at the ratio of 50% N. The reversal of charge point corresponding with point *a* must therefore lie at a higher $CaCl_2$ concentration (in Fig. 50 to the left of *a*).

At a ratio of 98% N, the excess of nucleate is even so large, that a mixed phosphatide + nucleate film is already formed at very low $CaCl_2$ concentrations. On increasing the $CaCl_2$ concentration the still phosphatide-rich film rapidly changes into a complete nucleate film (see the peculiar form of curve E in Fig. 49). Thus this colloid mixture shows only one reversal of charge point, viz. that of nucleate. Having given in Fig. 50 the course of the curve between a and b, we need not repeat the analogous explanation for the course between b and c.

Now turning to the question what significance these retrograde reversal of charge points have for the state of the colloids in the sol mixture itself, it will be clear that no interaction between these colloids is indicated by them. Indeed no signs of such an interaction manifest themselves, the sol mixtures being clear at all mixing ratios and $CaCl_2$ concentrations.

There remains the problem why such an interaction between the two colloids probably does not exist in mixtures, of phosphatide + nucleate + $CaCl_2$ and

---

[1] In § 2b (p. 277) we studied the rôle of colloid concentration on the formation of complete colloid films at the reversal of charge point of the colloid. As therefore we already used conditions most favourable to film formation, this second factor (the state of charge of the colloid) could not manifest itself. Certain irregular forms of U-log C curves (see p. 427, 428, Fig. 55 and 56) can be explained by this second factor.

why, as already mentioned, they do exist in mixtures of phosphatide + arabinate + La $(NO_3)_3$, the latter showing "normal" reversal of charge behaviour and also manifesting this interaction by the occurrence of flocculation.

One of the factors favouring electrical interaction is in general a low electrolyte content of the system, for interaction is generally decreased by salt (see p. 349, Ch. X, § 2 f). Indeed this factor may in part explain the above mentioned difference, for the reversal of charge concentrations of both colloids with La are smaller than with Ca. Therefore the mean salt concentration in the concentration range in which both colloids (here: phosphatide and arabinate) carry opposite electrical charges, is lower than in the combination phosphatide + nucleate + $CaCl_2$. Still the valency of the inorganic cation used is only one factor, others of more specific character also playing a rôle. For there exist cases in which two colloids interact in the presence of relatively large salt concentrations which will be discussed later (see Tricomplex Colloid Systems, p. 415, Ch. X, § 6). From that discussion follows that the conditions for interactions in the above example nucleate + phosphatide + $CaCl_2$ are indeed very unfavourable (see p. 421), both colloids belonging to the same class of colloids as regards the composition of the ionesed groups (phosphate colloids).

The absence of flocculation or coacervation phenomena in a mixture of oppositely charged colloids does not necessarily mean that colloid-colloid interaction is lacking. In mixtures of gum arabic and gelatin at lower pH values (e.g., pH 1.73 — 2.3) no coacervation occurs. Still interaction is present, for in Fig. 43 (p. 324) the curve continues its normal course in this pH range, absolutely different from that in Fig. 50 (p. 332), where the curve ends vertically at 100 % N. Thus in the sol mixture soluble combinations of gelatin and gum arabic are present and they are preferentially absorbed on the carborundum particles, which were used to measure the reversal of charge in these clear sol mixtures.

X. COMPLEX COLLOID SYSTEMS

H. G. BUNGENBERG DE JONG

Department of Medical Chemistry, University of Leyden.

§ 1. INTRODUCTION

a. Division into three principal types

In this Chapter we deviate from the division of the subject matter in this volume (II) of "Colloid Science" announced in Chapter I, § 5, p. 13 according to the kind of phase or state of dispersion of a multiphase colloid system. The possibility of the production of complex colloid systems is not connected with one of these modes of occurrence but with the electrolytic nature of the colloid(s) in question.

Under the term *complex colloid systems* are included colloid systems of all kinds, in which certain interactions, which we shall denote by *"complex relations"*, between the ions (in a broad sense) composing them are present. This concept cannot be settled here in a few words but will already be outlined more distinctly in § 2.

If one wants a provisional characterisation — which however leads one to expect too concrete ideas — one could say that complex systems are salt-like combinations either of colloids among themselves or of colloids and micro ions whereby one has not committed oneself to the character which the system may have in addition according to the classification principle chosen in Volume II (sols — colloid crystals — coacervates — flocculi — gels).

It is possible therefore to subdivide the complex systems according to the latter principle of classification into various cases but it does not strike the essential point which is characteristic of complex systems and indeed produces no rational order in the mutually very varied behaviour of the complex colloid systems.

After a division according to the number of colloids present into *"complex systems in the narrower sense"* (two colloid components) and *,,auto-complex systems"* (one colloid component) had appeared for some time to render good service, this division was also seen to be inadequate, when a kind of complex systems (p. 415, § 6) was found where this classification principle leads to great difficulties. Since then it has become clear that the rational division of the complex systems ought to rest on the type of ions between which the complex relations are enacted[1].

Three types of ions are known: cations, anions and amphoions. If the complex relations are present between ions of only one type — and these are then necessarily amphoions — we speak of *unicomplex systems;* if they are present between two types of ions — cations and anions — then we call them *dicomplex systems;* if on the other hand the complex relations are present between three types of ions at the same time

[1] H. G. BUNGENBERG DE JONG, *Proc. Koninkl. Nederland. Akad. Wetenschap.*, Amsterdam, 41 (1938) 776.

— amphoions + cations + anions — then this is expressed by the term *tricomplex systems*.

For the further characterisation of a complex system it is naturally useful to indicate also the nature of the system (for example, sol, coacervate, floccule, gel). This leads to a binary nomenclature in which first the principal type is given and then the nature of the system.

Thus a *unicomplex gel* is a system with the character of a gel the behaviour of which with regard to certain variables can be understood from the presence of complex relations between amphoions in this gel. Similarly a *dicomplex coacervate* is a coacervate in which complex relations are present between cations and anions. One can also denote the process by which such a complex system is produced by the same binary nomenclature. The last mentioned coacervate is the result of a *dicomplex coacervation*.

Similarly a *tricomplex flocculation* is a flocculation in which a disperse colloid-rich phase happens to separate out, in which latter phase complex relations are present simultaneously between amphoions, cations and anions.

b. Variants of the principal types

In the division of complex systems into three principal types it is expressly left indeterminate whether the ions in question are colloid ions or lie below the actually completely arbitrary limit and thus are "micro ions". Furthermore we here construe the term colloid ions very broadly and thus do not restrict ourselves only to ions of macromolecular colloids but also include in it the charged "particles" of association colloids. In practice it is however profitable to retain the above mentioned limit. According as the ions, which essentially take part in a complex system, lie above or below that limit, one can then foresee the following variants.

Unicomplex systems
1. Colloid amphoion + colloid amphoion
2. Colloid amphoion + micro amphoion
3. Micro amphoion + micro amphoion

Dicomplex systems
4. Colloid cation + colloid anion
5. Colloid cation + micro anion
6. Micro cation + colloid anion
7. Micro cation + micro anion

Tricomplex systems
8. Colloid amphoion + colloid cation + colloid anion
9. Colloid amphoion + colloid cation + micro anion
10. Colloid amphoion + micro cation + colloid anion
11. Colloid amphoion + micro cation + micro anion
12. Micro amphoion + colloid cation + colloid anion
13. Micro amphoion + colloid cation + micro anion
14. Micro amphoion + micro cation + colloid anion
15. Micro amphoion + micro cation + micro anion.

Of the 15 variants of the complex systems so obtained, there are three (underlined dotted) which contain no colloids and therefore must be considered as limiting cases.

The remaining 12 contain at least one colloid and thus form the real complex colloid systems. Only 6 of these (the underlined ones) have been realised and call for discussion in this chapter.

Now similar to the nomenclature in § 1a (p. 336) one can again go over to a binary nomenclature for the characterisation of a complex system, in which first the variant and then the nature of the system is mentioned. Of each of these variants one can thus foresee: sols, colloid crystals, coacervates, flocculi and gels.

Complex relations in colloid crystals have not been the subject of study and therefore will not be considered here.

By far the largest part of this chapter is devoted to complex coacervates and complex flocculations. Where possible a few words will be said in the appropriate sections on complex sols belonging to the same variant (p. 346, 392, 407, 413).

A few examples of complex gels belonging to variant 4 will be treated at the end of § 2. (p. 381, § 2u).

c. Order of the subject matter in relation to the role of the characteristic charge elements

In the discussion of the six known variants of complex coacervation or flocculation, we begin not with the unicomplex variant but with the dicomplex variants. On the one hand because the latter have been more fully studied and on the other hand because all predicable variants (No. 4, 5, 6, see Survey on p. 336) are also really known, as also the theoretically interesting limiting case (No. 7), "unmixing" (restricted solubility) in aqueous salt solutions. Our knowledge of unicomplex coavervation or flocculation, and similarly tricomplex coacervation or flocculation, is much more incomplete.

The limiting case No. 7, does not really belong to "Colloid Science" because both ions fall below the agreed dimensions; nevertheless No. 4, 5, 6 and 7 form a completely allied group. Characteristic properties of 4 are to be found not only in 5 and 6 where at any rate one colloid is present but also in 7 in which there is no longer any colloid at all. The limiting case No. 7 is therefore theoretically important because it shows us clearly that the essentials of the dicomplex colloid systems No. 4, 5 and 6 cannot be sought in the macromolecular structure of the colloid (or the association nature of the kinetic units in the case of the association colloids) but rather in the electrolyte nature of the macromolecule (or associate).

Indeed in the field of the complex colloid systems it is just the characteristic charge elements of the colloid (equivalent weight, polarisability of the ionised groups) and of the ions (valency, radius, polarisability) which play a decisive role. It is for this reason that these charge elements were already discussed in detail in chapter IX (p. 259).

In the discussion of dicomplex systems we do not begin with the apparently simplest case — the limiting case No. 7 — and by gradually replacing a microion by a colloid ion (No. 5 or 6) finally arrive at the apparently complicated case No. 4 where both ions are colloid ions. Such a treatment is certainly not the way because, going from No. 7 via No. 5 and 6 to No. 4 does not in reality mean a complication but rather a simplification.

Variant No. 4 is in fact the simplest case, indeed in this case the complex relations occur between a large organic cation (colloid cation) and a large organic anion (colloid anion).

As we have seen in the preceding Chapter (p. 300, Ch. IX, § 4) this involves
a. the true reversal of charge concentrations being very small;
b. polarisation phenomena in the interaction of the cation and anion being absent.

In the variants No. 5 and 6 (*a*) need not hold in general, which brings with it complications. Also (*b*) no longer holds in variant 6 when the micro cations are small. Extra complications then occur because of this, that is to say specific differences begin to manifest themselves according to whether the colloid is a sulphate-, a carboxyl- or a phosphate colloid.

Arriving finally at No. 7 we are entirely in the region in which all sorts of as yet inestimable specific factors govern the occurrence or otherwise of an ion "unmixing".

After the dicomplex systems the unicomplex ones will be cursorily discussed and finally we shall deal with the tricomplex systems.

In the two known variants of the latter specific factors, as we shall see, play an extraordinarily prominent role.

§ 2. DICOMPLEX SYSTEMS I.
VARIANT COLLOID CATION + COLLOID ANION

a. Introduction

Protein cations appear throughout this section as the colloid cations. On account of the ampholytic nature of the proteins, there is really no question of an exclusively positively charged protein cation at a pH little lower than the isoelectric point but rather there is a smaller number of negatively ionised groups as well as the positively ionised ones. According to experience a protein in the variant colloid cation + colloid anion is active with the excess of the positive charges over the negative, thus with the algebraic sum of its charges as it were.

This excess is a function of the pH. It is zero at the isoelectric point and consequently the equivalent weight of the protein as protein cation is infinitely large. On gradual lowering of the pH whereby this excess is first produced and at first continually increases, the equivalent weight of the protein cation therefore gradually decreases in the above sense (see p. 273, Ch. IX, § 1f).

It may also be remarked that in the unicomplex and tricomplex systems a protein is *not* active with the excess of the one kind of charge over the other but with them what matters is the electric multipole character of the protein. A too large excess is with them even a hinderence for the occurrence of the systems in question.

b. Complex coacervation in the narrow sense

In the older literature it has already been mentioned that with a weakly alkaline reaction (dilute ammonia) solutions of protamines or histones — which are "basic proteins", that is to say, proteins the isoelectric point of which is $>$ 7 — mixed with solutions of "acid proteins" — which are proteins whose isoelectric point is $<$ 7 — give precipitates which separate out either as flocculi, or as sticky masses or as liquid "oil-like" drops[1].

In many cases it is indicated that the mixing proportion of the sols is important

---

[1] Compare for example, A. HUNTER, *Z. physiol. Chem.* 53 (1907) 526 for precipitation of the protamine clupein with other proteins.

and that excess of one of the components can cause the precipitate to dissolve.

Here and there it is also observed that these precipitates dissolve in dilute salt solutions. In general these precipitates have been considered as salt-like protein — protein compounds. Expressed in modern terms the condition for the production of these precipitates is that at any rate the pH, at which the experiments are made, lies between the two isoelectric points so that the one protein is present as cation, the other as anion. In addition the mixing proportion must not be too unfavourable.

The phenomenon is not restricted to the combination protein — protein; but can also occur when one of the colloids is a substance of the acidoid type, provided again that the pH is such that the protein is positively charged, that is to say at a pH < I.E.P.

The oldest observations in this field concern combinations of basic proteins[1] with nucleic acids, from which the representation of the nucleoproteins as salt-like compounds resulted.

In many of these cases the precipitate again had glutinous properties, or separated out as "oil-like" drops[2]. Since then the number of cases in which a colloid-rich liquid separation, that is to say, coacervation, was observed on mixing oppositely charged hydrophilic sols has increased very greatly.

To distinguish this from other types of coacervation (p. 250 and 255, Ch. VIII, § 5 and 8), this was called *complex coacervation*. The prefix "complex" is hereby meant to express that the two colloids which form the coacervate together with water, have united as a result of a contrast of charge.

One can in fact also have the case that a coacervate contains two colloids because each of these colloids, if present alone, forms a coacervate under the conditions of the experiment and these two coacervates are mutually miscible. Such a coacervate can be called a *mixed coacervate*. In § 2 t we shall even encounter examples in which two complex coacervates form a mixed coacervate containing three colloids (see p. 378).

Though newer developments allow to use the term "complex coacervation" also in a much wider sense (see pag. 337) than that given above, we will for brevity's sake retain its original meaning throughout this whole § 2 and will therefore omit the restriction: belonging to the variant 4 of the survey on p. 336.

c. Reversibility of complex coacervation

If one adds a little HCl to a warm mixture of dilute gelatin and gum arabic sols, complex coacervation first takes place at a certain pH value below 4.8 (the I.E.P. of the gelatin used). The liquid becomes turbid and the presence of coacervate drops can be readily detected microscopically[3]. If now one adds some NaOH to a pH above 4.8, the coacervation disappears again, to reoccur again after acidification. This can be repeated a number of times. It can however happen that with increasing number of repititions of alternate additions of NaOH and HCl, the turbidity becomes steadily weaker and finally does not occur at all. This is explained by the formation of an increasing amount of NaCl. It is in fact characteristic of complex coacervation

---

[1] A. KOSSEL, *The Protamines and Histones*, 1928. Longmans, Green and Co., London, New York, Toronto, see pp. 27, 31.

[2] For example H. STEUDEL and E. PEISER, *Z. physiol. Chem.*, 122 (1922) 298.

[3] The coacervate drops coming in contact with the glass surface of the slide soon become invisible as a result of wetting. To prevent this starched microscope slides may be used (see p. 435).

that this is hindered by indifferent salts at sufficiently high salt concentrations or if we already have a coacervated system, the coacervation is suppressed by the addition of an indifferent salt (see in more detail below, p. 349, § 2f).

The described reversibility of complex coacervation is not restricted to the combination gelatin — gum arabic but holds generally, provided no secondary changes of a different kind occur.

Typical coacervation occurs with the combination serum albumin — gum arabic on acidification. With the microscope one sees beautiful thoroughly liquid coacervate drops, which readily coalesce with each other (see p. 233, Fig. 2). Immediately after the production of the coacervation the latter is completely reversible: added KCl completely suppresses the coacervation. If however one adds KCl some time after the coacervation has been brought about, the system no longer clears completely but remains turbid to an increasing extent the longer one has waited before adding KCl.

The cause of this is the denaturation of the serum albumin in the complex coacervate. The complex relations between gum arabic and serum albumin are indeed suppressed by the indifferent salt whereby the first and not yet denatured part of the serum albumin goes into solution, the denatured fraction of the serum albumin remains behind as insoluble matter.

Another change of a secondary kind whereby the suppression with indifferent salts is destroyed can be observed in the complex combinations in which gelatin is one of the colloid components. These combinations behave in a completely reversible manner at higher temperatures (e.g. 40°). If one allows the coacervated system to cool, the coacervate drops become solid (gelation) and adding of indifferent salts no longer causes them to go into solution, although they can indeed swell considerably and can thereby become less clearly visible (see for further information p. 381 § 2 u and p. 450, Fig. 18).

These changes of the coacervate are based entirely on the gelating power of the gelatin component and since this gelation is reversible they are again completely suppressed by warming. The gelated coacervate drops again become liquid and they now again go completely into solution on addition of indifferent salts.

d. Rôle of pH and mixing proportion in complex coacervation

The combination gelatin — gum arabic can be considered as the most favourable object as yet for the study of coacervation. In this case the complex coacervate has relatively little viscosity[1] and consequently readily fuses to a single transparent liquid layer whereby it becomes possible to take samples of coacervate layer and equilibrium liquid and investigate them as regards their composition. The two colloids can be kept in the dry state for unlimited times and show no denaturation phenomena in solution. The only factor to which one must pay attention is the temperature, since one otherwise obtains the complications mentioned above as a result of gelation.

That this factor was unknown is certainly a contributary cause that TIEBACKX's [2] investigation on this combination gelatin — gum arabic has contributed so little to a clearer view of the problem.

---

[1] A single measurement, by which the validity of Poiseuille's law was proved, furnished a value of 20x that of water at the same temperature (see note 3 on p. 245).
[2] F. W. TIEBACKX, *Kolloid Z.*, 8 (1911) 198; 9 (1911) 61; 21 (1922) 102.

When one works at a temperature above about 33° one is completely safe. Furthermore this complex coacervation, according to volume measurements and analyses is practically independent of the temperature[1] in the range 33—50°.

BUNGENBERG DE JONG and DEKKER[2] investigated this combination at 40° and employed for the purpose various methods besides analyses (p. 355 ff.) such as measurements of viscosity, of turbidity, of coacervate volume and of electrophoresis.

First of all a summary of the variables may be given:

Complex coacervation depends on:
1. the pH.
2. the mixing proportions of the isohydric sols.
3. the initial concentration of the isohydric sols.
4. the possible presence of indifferent salts.

1. No complex coacervation occurs at a pH higher than the isoelectric point of the gelatin used (4.8) however one changes the other factors.

2. On account of the important role of the pH which regulates the (apparent) equivalent weight of the gelatin and of the gum arabic (dissociation of the arabinic acid) the appropriate way is to bring the gelatin and gum arabic sols to the same pH with HCl before mixing them. It has been found that the pH of such isohydric sols does not change on mixing and that when coacervation occurs the pH of coacervate and equilibrium liquid is the same.

Fig. 1 refers to the changes of viscosity which occur in a number of isohydric series of mixtures as a function of the mixing proportions of 0.67% sols. The mixing proportion expressed in % of the gum arabic sol is the abscissa, while as ordinate one takes the experimentally determined quantity $\frac{\eta_s - \eta_o}{\eta_o}$ expressed in % of the

value $\frac{\eta_s - \eta_o}{\eta_o}$ calculated from the corresponding expressions for the two isohydric sols assuming additivity.

The figure shows that the behaviour of the viscosity in mixtures of isohydric sols is additive at pH 5.06 and pH 1.22, that is to say at pH values at which no complex coacervation occurs.

At the other pH values there are indeed deviations from the additive behaviour. In this pH range the gelatin is positively and the gum arabic still negatively charged[3].

The deviations from additivity are obviously very closely connected with the interaction between the positive gelatin and the negative gum arabic. If this interaction is not powerful enough the sol mixtures are clear even at the most favourable mixing proportion (e.g. the series at pH 2.06). With sufficient interaction complex coacervation occurs in a range of mixing proportions round about the minimum of the curves.

If one plots the position of the minimum expressed in % A as a function of

[1] H. G. BUNGENBERG DE JONG, E. G. HOSKAM and L. H. v. D. BRANDHOF-SCHAEGEN, Proc. Koninkl. Nederland. Akad. Wetenschap., Amsterdam, 44 (1941) 1104.
[2] H. G. BUNGENBERG DE JONG and W. A. L. DEKKER, Kolloidchem. Beih., 43 (1935) 143; 43 (1936) 213.
[3] The electrophoretic velocity of gum arabic decreases on decrease of pH and reaches a value of practically zero at pH 1.7 (see p. 325, Fig. 44).

the pH, then there results a curve (see Fig. 2a) which shows the position of opti-
mal interaction as function of pH and mixing proportion of the colloids.
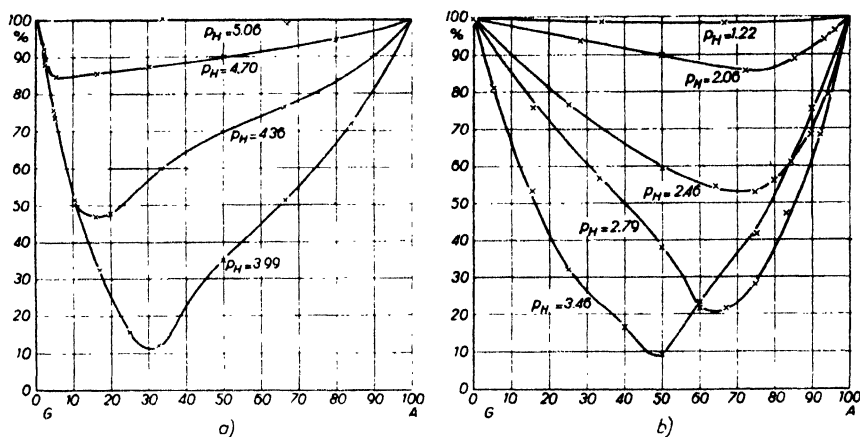
Fig. 1. Viscosimetric investigation of complex coacervation in mixtures of isohydric 0.67% sols
of gelatin (G) and gum arabic (A).

Ordinates: the experimentally determined quantity $\dfrac{\eta_s - \eta_o}{\eta_o}$ expressed as a percentage of the value
$\dfrac{\eta_s - \eta_o}{\eta_o}$ calculated from the corresponding expressions for the two separate isohydric sols assuming
additivity. The ordinate values for these latter sols are thus always 100% independently of the pH
of the series of mixtures. Values lower than 100 % denote colloid-colloid interaction.
Abscissae: mixing proportion of the two isohydric sols expressed in % of the gum arabic sol.
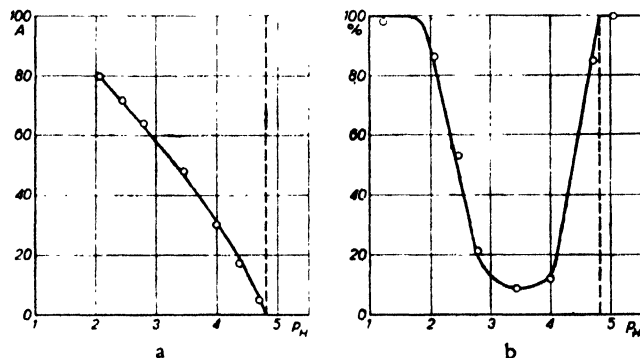
Fig. 2. Abscissa values (left graph) or ordinate values (right graph)
of the minima in Fig. 1 plotted against the pH of the isohydric
series of mixtures (see text).
The vertical dotted line marks the isoelectric point of the gelatin used.

If one plots the
ordinate values of the
minima in Fig. 1 as a
function of the pH, there
results a curve with a
minimum at about pH
3.5 (see Fig. 2b), that is
to say, within the range
of pH, within which gel-
atin and gum arabic are
oppositely charged (pH
4.8—1.7), there is also a
value of the pH which is
most favourable of all for
complex coacervation.

In the case also when one studies complex coacervation by means of turbidity
measurements (a method which can be employed at smaller colloid concentrations)
one arrives at similar results. Compare Fig. 3 and 4, which refer to isohydric series
of mixtures of 0.05% sols.

These and other measurements (also with much more dilute sols) lead to the conclusion that the complex coacervation depends on a union of gelatin cations and arabinate anions, which within each isohydric series of mixtures is possible only optimally at one particular mixing proportion. At fairly high values of pH where the
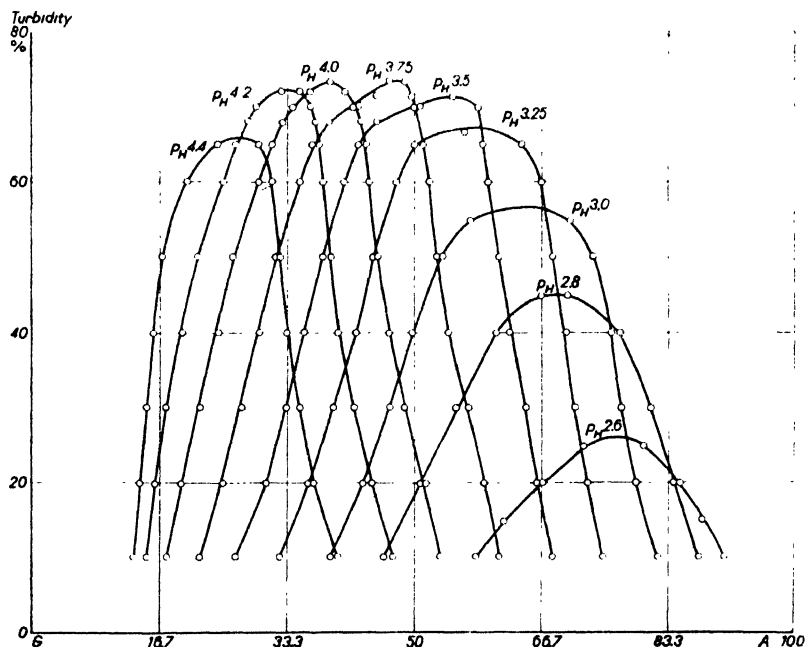
Fig. 3. Turbidity curves for isohydric mixtures of 0.05% gelatin and gum arabic sols.
Ordinates: turbidity at 42°, 10 minutes after mixing the sols[1].
Abscissae: mixing proportion of the isohydric sols expressed in % of the gum arabic sol.
The curves have been deduced from Fig. 5 by reading off the mixing proportions at a number of constant pH's at which 10, 20 .... etc. % turbidity occurs.

(apparent) equivalent weight of the gelatin is still fairly high, the equivalent weight of the gum arabic still low, only little gum arabic is necessary for much gelatin for the mutually equivalent union of gelatin cations and gum arabic anions. That is to say, the optimum coacervate lies at a low value of the mixing proportion expressed in % A.

According as the pH is chosen lower, the (apparent) equivalent weight of the gelatin decreases and that of the gum arabic increases (decrease of the dissociation of the COOH group) and thus the mixing proportion of the optimum coacervation expressed in % A also increases.

The curves in Fig. 2a and 4a must thus begin at the isoelectric point of the gelatin (at this point the apparent equivalent weight of the gelatin is infinite and

---

[1] Turbidity was measured here and also in the figures 4, 11, 16, 37, 40-44, 45, 49, 53, 54 and 57-60 with a extinctometer of Moll using a layer of thickness 10 mm. As measure of turbidity the values $100 \dfrac{I_0 - I}{I_0}$ are taken, in which $I_0 =$ the intensity of the incident light and $I =$ the intensity of the light after passage of the cuvette contents.
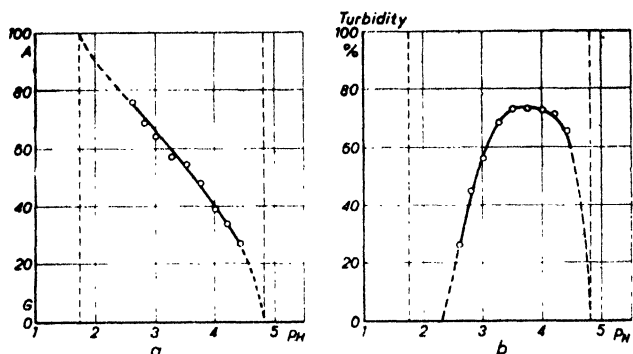
Fig. 4. Abscissa values (left graph) or ordinate values (right graph) of the maxima in Fig. 3, plotted against the pH of the isohydric series of mixtures (see text).
The vertical dotted line marks the isoelectric point of the gelatin used.

thus the optimum mixing proportion expressed in % A is just zero). Similarly the curves must end at the discharge point of the gum arabic (pH 1.7), at an optimum mixing proportion of 100% A.

However, although there is a corresponding definite equivalent mixing proportion at each pH in the pH range in question (4.8—1.7) in which gelatin and gum arabic are oppositely charged, that is
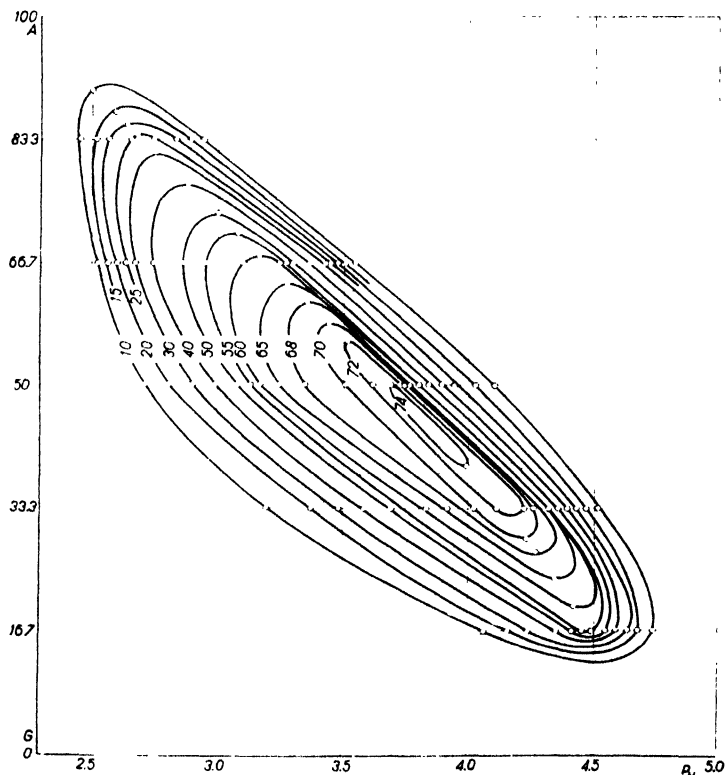
Fig. 5. Curves of equal turbidity in the mixing proportion-pH diagram for complex coavervation of 0.05% gelatin and 0.05% gum arabic sols.
At some five constant mixing proportions the turbidity had been determined as a function of pH and the pH values read off from this at which 10—20 .... etc. % turbidity occurs.

not to say that the coacervation will thereby be equally vigorous. On the contrary the coacervation is only vigorous at pH values which lie not too close to the extremities of this pH range (see Fig. 5). At pH 4.7 the coacervation is only very slight and at pH 2.06 it is even absent.

When we continue to cling to the nature of complex coacervation as a process which is allied to salt formation between gelatin cations and arabinate anions, it will be clear that not one single equivalently constituted colloid-colloid salt of this kind exists but a whole series whose composition depends on the pH. (This is further confirmed in the discussion of the analytical composition of the complex coacervates. See p. 359, § 2i).

They are obviously of different solubility. Now it is probable that then the smallest solubility will be present when as large a number as possible of "salt bonds" per unit weight of gelatin-arabinate is present. For everything points to this salt bond being the immediate cause of the complex coacervation; indeed gelatin and gum arabic in the uncharged state are both soluble.

In one gram of the optimum gelatin-arabinate at pH 4.7 only a relatively small number of salt bonds is present; indeed the gelatin-arabinate then consists of 95% gelatin with a very high equivalent weight and 5% arabinic acid.

As the pH decreases the number of salt bonds per gram of gelatin-arabinate increases, goes through a maximum and then again decreases. This therefore explains the existence of a most favourable pH for complex coacervation: for this there would follow from Fig. 4b a pH = about 3.7 for 0.05% sols, while from Fig. 2b for 0,67% sols a somewhat lower value (pH = about 3.5) appears to follow.

e. Behaviour of complex coacervate drops in an electric field

If one introduces a coacervated system, consisting of coacervate drops suspended in their equilibrium liquid, into a d.c. electric field three phenomena [1] can be observed simultaneously: 1. electrophoresis; 2. deformation; 3. disintegration phenomena

1. The *electrophoric sign* of the charge varies in an isohydric series of mixtures (see p. 321, Chapter IX, § 6a). With mixing proportions which are poorer in gum arabic than the optimum the drops are positively charged, with mixing proportions richer in gum arabic than the optimum they are negatively charged. Reversal of charge therefore occurs at or very near to the optimum mixing proportion. Compare Fig. 6 which refers to isohydric series of mixtures of 0.05%.

In each series of mixtures the position of the optimum turbidity was determined (black dots) while at the same time the position of the reversal of charge point (open circles), was determined for a number of mixtures (see p. 322, Ch. IX, § 6b).

We conclude from Fig. 6 that optimum coacervation and reversal of charge coincide very closely. The various points fit with the view of complex coacervation as the separation of a colloid-colloid salt, which at the reversal of charge point is composed of equivalent (or nearly equivalent) quantities of gelatin cations and arabinate anions (see p. 326, Ch. IX, § 6c).

If the sol mixture contains too much gelatin the coacervate is positively charged, if it contains too much gum arabic it is negatively charged. The change in composition of

---

[1] H. G. BUNGENBERG DE JONG and W. A. L. DEKKER, *Kolloid Beih.* **43** (1935) 331; *Biochem. Z.*, **221** (1930) 403.

the coacervate in an isohydric series is further dicusssed in § 2g (p. 355) and § 2j (p. 360).
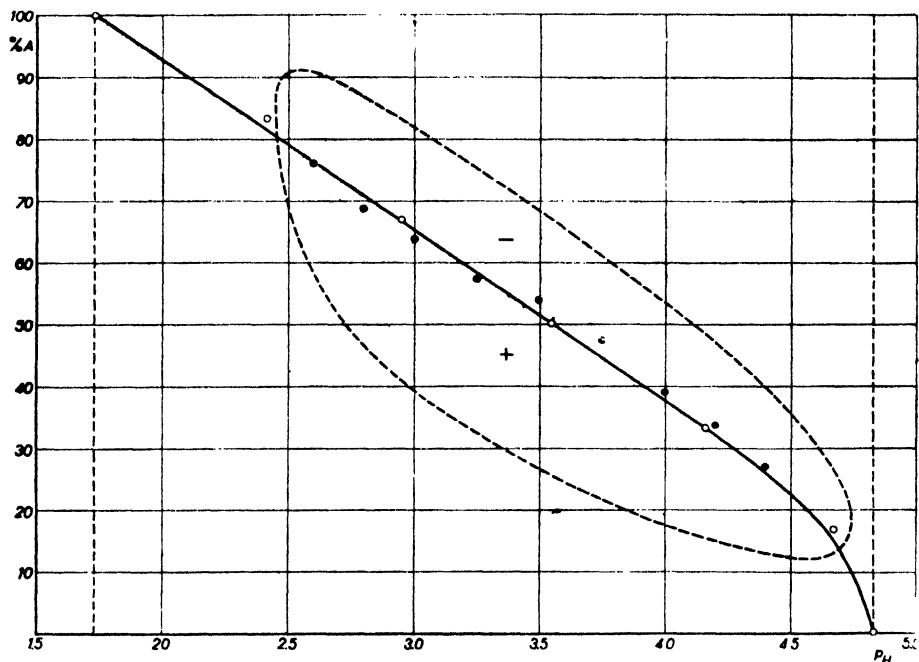
Fig. 6.   Coacervation region, optimum coacervation and reversal of charge in the mixing proportion⁻
pH diagram (0.05% sols).
Ordinates: Mixing proportions in % of gum arabic sol.
Abscissae: pH.
O : reversal of charge points.
● : optimum coacervation within the isohydric series of mixtures.
vertical dotted lines: position of the isoelectric point of the gelatin (right) and position of the pH
at which the gumarabic no longer has any charge electrophoretically.
closed dotted curve: curve of 10% turbidity.
The actual coacervation region extends a little outside this curve on all sides. This latter region
is surrounded by homogeneous sol mixtures. Those which lie close to the coacervation region may be
considered as *complex sols*, i.e. sol mixtures in which complex relations are present, but which
latter are to weak to cause actual separation into two phases.

In Fig. 6 is also drawn a closed curve (dotted) on which the turbidity amounts
to 10% (the outermost curve of Fig. 5) that is to say, where the complex coacervation
is only very weak.
The region in which complex coacervation occurs in the mixing proportion —
pH diagram, has thus the same shape in principle, but is somewhat more extensive
in all directions than the area enclosed by the dotted curve.
The close connection between complex coacervation and reversal of charge is
here clearly visible, indeed this coacervation field in the diagram is intersected by
the full drawn curve (= reversal of charge), which latter divides the field into two
halves, in which the coacervate drops have opposite charges.

2. The *deformation of coacervate drops* also occurring in a d.c. field, can be observed undisturbed in an a.c. field (for example of frequency 50) since electrophoresis and disintegration phenomena are hereby excluded.

The coacervate drops are thereby flattened in a direction perpendicular to the lines of force of the field (see Fig. 7). On switching of the field they return to the

Fig. 7. Deformation of complex coacervate drops in an a.c. electric field.
Arrow: direction of the lines of force.
Deformation to almost flat discs takes place only with a sufficiently powerful field. With less powerful fields the drops assume an ellipsoidal shape and the degree of deformation is still greatly dependent of the original diameter of the drops (small drops deform less readily than large).
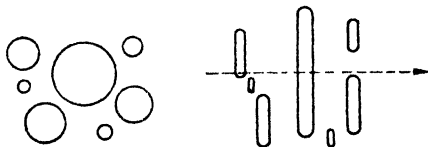
spherical shape. As far as it has been investigated this flattening appears to be closely connected with a phenomenon described by BÜCHNER[1] and then must depend on a difference in conductivity of coacervate and equilibrium liquid. If the former is smaller than the latter the flattening must take place perpendicular to the lines of force, in the opposite case an extension in the direction of the lines of force must take place. Measurement shows that the coacervate has a smaller conductivity than the equilibrium liquid so that the observed flattening perpendicular to the lines of force is in agreement with expectation.

The behaviour of vacuoles, which are enclosed in the complex coacervate drops, is also in agreement with this explanation. The contents of the vacuole are equilibrium liquid and thus have a greater conductivity than the coacervate surrounding them. In accordance with expectation the vacuole is stretched to a spindle-shaped body in the direction of the lines of force while the coacervate drop in which this vacuole is embedded is flattened in the usual way.

3. In a d.c. field changes occur in the interior of the complex coacervate drops which, if the field is strong enough, soon lead to the disappearance of the large coacervate drops and cause the production of a large number of smaller ones.

We shall omit any detailed description of streaming and vacuolation processes in the coacervate drop. They are discussed in Chapter XI (see p. 452).

Fig. 8 reproduces schematically some salient phenomena. From this figure it appears that the phenomena with sufficiently negatively charged coacervate drops are the mirror image of those with sufficiently positively charged drops. (The phenomena are more complicated in the neighbourhood of the reversal of charge point).

In these schematic figures the small vacuoles initially occurring and their transport by streamings are left out of consideration. After some time small vacuoles unite into larger ones, while fresh small vacuoles continue to be formed possibly at a reduced rate.

The schematic figure now indicates the fate of the large vacuoles. The latter are transported in the coacervate to the cathode in the case of positive coacervate drops, to the anode however in the case of negative coacervate drops. This transport is probably an electrophoresis, in fact not only vacuoles but also all inclusions (organic liquid drops, solid particles) move in the same sense (p. 445, Fig. 12). These inclusions behave as if

---

[1] E. H. BÜCHNER and A. H. H. VAN ROYEN, *Kolloid Z.*, 49 (1929) 249.

they have the same sign of charge as the coacervate in which they are enclosed[1].

The vacuoles arriving at the edge of the coacervate drop protrude there and soon the coacervate lamella lying in between bursts. The vacuole contents are thereby ejected into the surrounding liquid, with which they are seen to be completely miscible and the coacervate drop rounds itself off on that side.

At the other side of the coacervate drops many new small coacervate drops form in the adjacent equilibrium liquid at a short distance from the coacervate surface.

This phenomenon is also the mirror image with positive coacervate drops of that with negative drops.

A simple explanation of these "*disintegration phenomena*" is obtained by assuming that the two colloid components in the complex coacervate are not really bound into a rigid salt but that the gelatin cations and the arabinate anions are displaceable in an electric field. The gelatin cations will move in the direction of the cathode, the arabinate anions in the direction of the anode. If further we assume that these colloid ions or at least the colloid ions which are oppositely charged from the coacervate surface can also pass through the coacervate surface, then we can predict the place where new coacervate drops are produced in the equilibrium liquid.

If we start from a positively charged coacervate drop, this signifies that an excess of gelatin is present in the total system. According to analyses this excess is present partly in the equilibrium liquid. If one adds a little isohydric gum arabic to this equilibrium liquid coacervation sets in. Now we see that in an electric field new coacervate drops are produced at the anode side of the drops. That is to say, at this side of the coacervate drop arabinate ions originating from the coacervate are transported into the equilibrium liquid.

If gelatin cations can possibly leave the positively charged coacervate and that

Coacervate drops positive   Coacervate drops negative
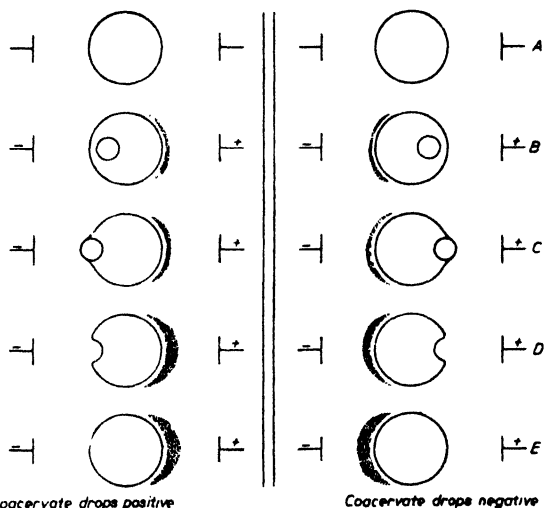
Fig. 8. Some points of detail in the behaviour of complex coacervate drops in a d.c. electric field (schematic).
Only two points of detail are depicted schematically here. 1st. The larger vacuoles formed some time after the application of the field are displaced towards one side of the coacervate drop and there expel their contents outwards. 2nd. On the other side a large number of new small coacervate drops are formed outside the coacervate drop.
The events with negative coacervate drops are the mirror image of those with positive drops.

will then be on the side of the cathode, they will in any case not give rise to the formation of new coacervate drops. Indeed the equilibrium liquid already contains gelatin in excess and addition of a little isohydric gelatin sol to it does not cause coacervation.

[1] H. G. BUNGENBERG DE JONG and A. J. W. KAAS, *Bioch. Z.* 232 (1931) 338.

Thus this explains why positive coacervate drops get a crown of new coacervate drops in the equilibrium liquid exclusively on the anode side.

Mutatis mutandis the production of new coacervate drops on the cathode side follows for negative coacervate drops, indeed the gelatin cations leaving on this side because of the electric field pour into an equilibrium liquid which contains an excess of arabinate anions; a situation which leads to fresh formation of coacervate drops.

f. Salt Influences on Complex Coacervation

Complex coacervation is hindered when a sufficiently high concentration of indifferent salt is present; similarly an already existing coacervation is suppressed by addition of salt. The salt concentration required for the suppression is not the same for different salts.

If one compares salts with a monovalent anion or monovalent cation, then the suppressive action is seen to increase with the valency of the other ion. The suppressive action then decreases (that is to say, the salt concentration required increases from left to right) according to the sequences:

$4 — 1 > 3 — 1 > 2 — 1 > 1 — 1$: *valency rule of the cations*
$1 — 4 > 1 — 3 > 1 — 2 > 1 — 1$: *valency rule of the anions.*
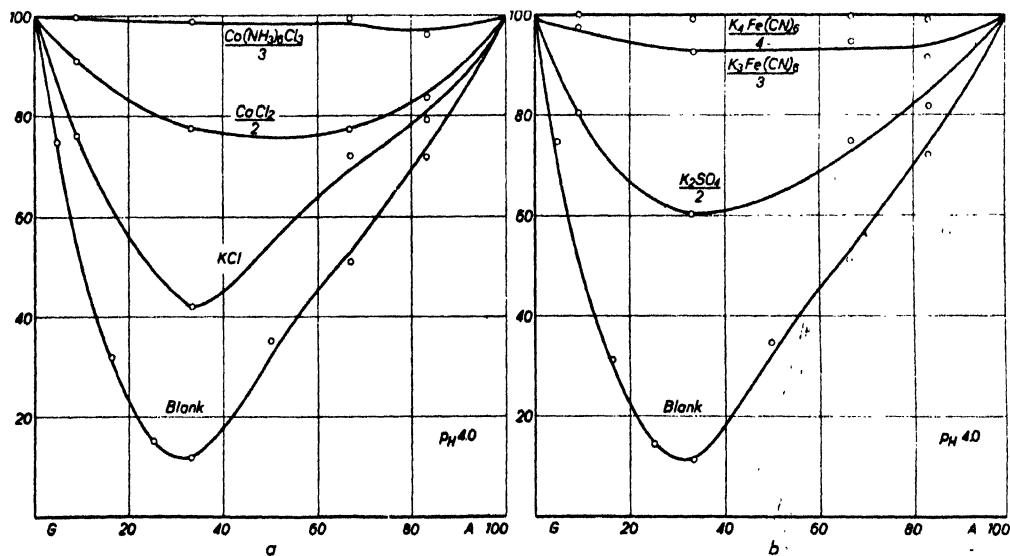
Fig. 9. Viscosimetric investigation of the suppressive action of indifferent salts on the complex coacervation of 0.67% gelatin and gum arabic sols at pH = 4.0.
Ordinates and abscissae as in Fig. 1 (p. 342); the lowest curve refers to sols without added salt; the remaining curves to sols which contain 13.3 m. eq. per 1. of the salt indicated.
The very large deviation from the additive behaviour in the blank, decreases through the added salt and the more strongly the more the valency of the cation (left) or of the anion (right) increases. Parallel with the return to additive behaviour the turbidity caused by the coacervation also decreases (round about the minimum the mixtures with KCl and CaCl$_2$ are still turbid but with Co(NH$_3$)$_6$Cl$_3$ all are clear; similarly K$_2$SO$_4$ is still turbid, K$_3$Fe(CN)$_6$ opalescent and K$_4$Fe(CN)$_6$ clear).

We can collect these two rules into the *"double valency rule"*, to distinguish it from another valency rule given below (p. 352-353).

Fig. 9 gives an example where the return to additive behaviour of $\dfrac{\eta_s - \eta_0}{\eta_0}$ is employed as a measure of the suppressive action.

Fig. 10 also gives an example of the cation valency rule in which the return to additive behaviour of $\dfrac{\eta_s - \eta_0}{\eta_0}$ is measured at a constant mixing proportion of the colloids. One sees that added cane sugar (non-electrolyte) in the concentrations in question here has no influence on the complex coacervation, but the four salts do have such an influence. They reduce the deviations from the additive behaviour and indeed at smaller concentrations the greater the valency of the cation.
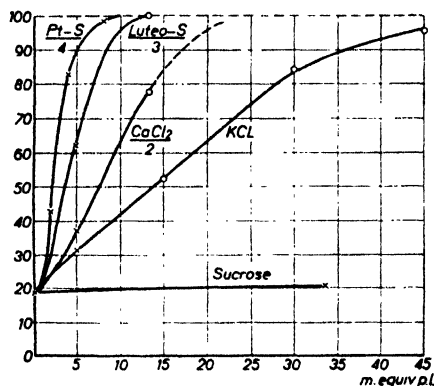
Fig. 10.   Return to additive behaviour of $\dfrac{\eta_s - \eta_0}{\eta_0}$ through added salts of the type 1—1, 2—1, 3—1, 4—1 and inactivity of sucrose in this respect.
Ordinates: as in Fig. 1 and 9.
Abscissae: concentration of the salts in m. eq. per 1. and of the sucrose in millimol per 1.
Pt-S   abbreviation for [Pt(en)$_3$] (NO$_3$)$_4$
Luteo-S   ,,   ,, [Co(NH$_3$)$_6$] Cl$_3$.

Fig. 11 gives another example in which we see the turbidity caused by complex coacervation in very dilute sol mixtures (0.01%) decrease and finally disappear altogether through added salts where the valency rule of the cations appears from a, that of the anions from b.
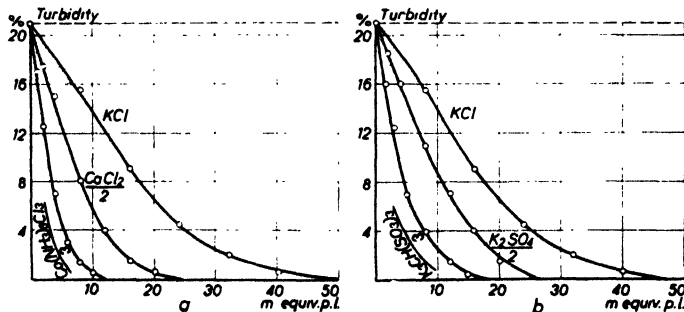
Fig. 11.   Double valency rule in the suppression of complex coacervation of 0.01% gelatin and gum arabic sols.
Mixing proportion of the colloids: 50% A, pH = 3.41.
Ordinates: turbidity 15 min. after mixing (42°).
Abscissae: salt concentration in m. eq. per 1.
a: valency rule of the cations.
b: ,,   ,,   ,, ,, anions.

Fig. 12 gives similar results with the so-called coacervate volume method. In contrast to the viscosimetric or turbidity method this is a direct method. It is serviceable in sufficiently concentrated colloid mixtures. In it one reads off directly the volume of the separated coacervate in a graduated tube (after sedimentation and coalescence to a homogeneous liquid column). Here also we see that added salts suppress coacervation and that the double valency rule holds.
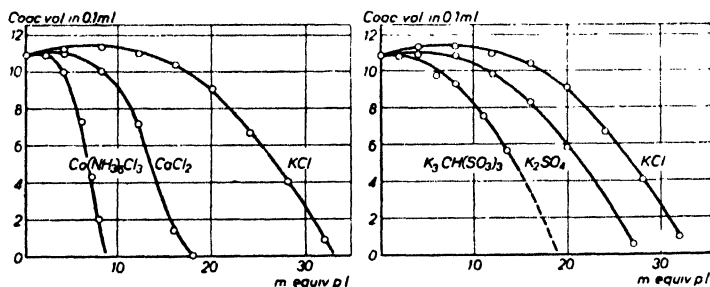
Fig. 12. Double valency rule in the suppression of the complex coacervate gelatin — gum arabic (coacervate volume method [1]).
Ordinates: coacervate volume in 0.1 cc produced from 25 cc sol mixture (2% air dry, mixing proportion 55% A, pH = 3.7).
Abscissae: salt concentration in m. eq. per l.

If one regards the complex coacervate as a slightly soluble compound of polyvalent colloid ions, one can understand the occurrence of the double valency rule as increase of solubility as a result of the shielding of the colloid cations by the anions of the added salt and of the colloid anions by the cations of the added salt.

Since this shielding increases with the valency of the shielding ion, the double valency rule follows from this. Indeed in it we always keep the valency of one ion constant (the anion in the valency rule of the cations; conversely the cation in that of the anions) and increase the valency of the other ion [2].

The required concentration for one chosen salt just to be able to suppress coacervation can be called the *salt resistance*. Thus we can speak, for example, of KCl resistance. It now appears that this KCl resistance for the coacervates is not constant in an isohydric series of mixtures but reaches a maximum in the neighbourhood of the optimum mixing proportion (that is to say, near the reversal of charge point) while this salt resistance decreases on either side. Compare Fig. 13 in which the arrows indicate the reversal of charge points of the series of mixtures in the absence of added salt.

The maximum salt resistances of isohydric series of mixtures are also not equal among themselves but these reach a highest value at a pH value which lies in between the isoelectric point of the gelatin and the discharge point of the gum arabic (in Fig. 13, at about pH 3.5—3.7).

Finally the salt resistance also depends on the colloid concentration of the two

---

[1] H. G. BUNGENBERG DE JONG and C. VAN DER MEER, Proc. Koninkl. Nederland. Akad. Wetenschap., Amsterdam, 45 (1942) 490.
[2] Specific differences between ions of the same valency are discussed later. See p. 376, § 2 s.

isohydric sols. It decreases on increase of these concentrations and even becomes zero at certain colloid concentrations, that is to say, mixing of the two isohydric sols no longer gives any complex coacervation although the pH is favourable. This is the case with gelatin—gum arabic at pH 3.5 if the colloid concentrations of the sols are higher than 5%. The decrease of the salt resistance and its becoming zero on increase of the colloid concentration can be understood after reading p. § 2m (p. 366), from which it appears that in the coacervation itself a neutral salt is already formed and the concentration of it increases with the colloid concentrations.
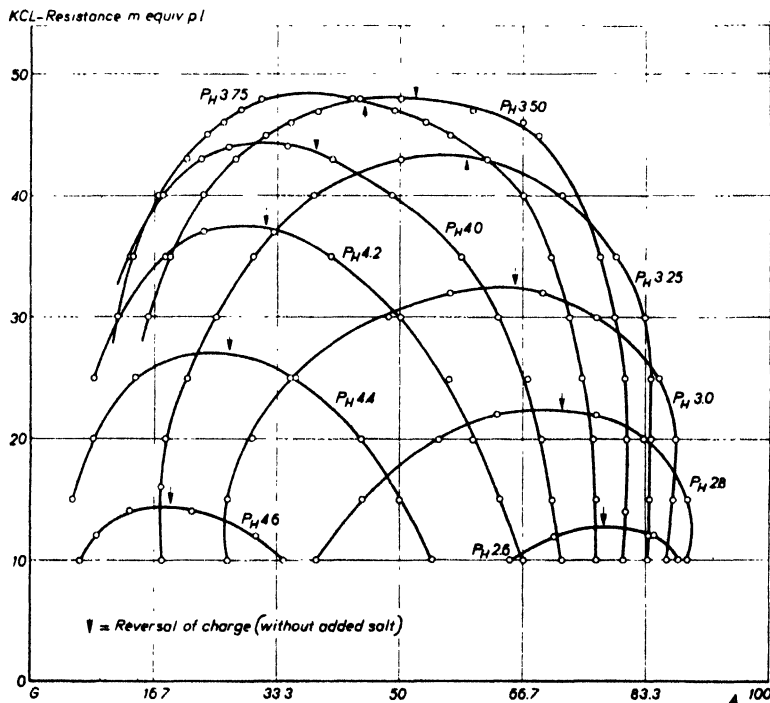
Fig. 13.   KCl resistance of the complex coacervates as a function of the mixing proportion of the isohydric 0.05% gelatin and gum arabic sols.
Since it is difficult to indicate just where the turbidity becomes zero, one here takes as the KCl resistance (ordinates) the KCl concentration (in m.eq. per l.) at which the turbidity is lowered to a very small value (2%).
Abscissae: Mixing proportion of the isohydric sols expressed in % of the gum arabic sol.

We now turn to another salt rule, the "*continuous valency rule*", in which the same salt symbols occur as in the double valency rule, but in which they are arranged in another way.
It was first observed in the influence of salts on the electrophoretic velocity of complex coacervate drops and complex flocculi[1].

[1] H. G. BUNGENBERG DE JONG and J. L. L. F. HARTKAMP, *Rec. Trav. Chim.*, 53 (1934) 622.

As already stated these coacervate drops can be negatively charged, uncharged or positively charged according to the mixing proportion of the colloids and the pH.

If now one investigates the influence of salts of the same valency types as occur in the "double valency rule", the curves for these salts are seen to form a fan in which the salts are always arranged in the same sequence; Fig. 14 gives an example for a positively charged coacervate system, but the sequence is the same for an uncharged or a negatively charged system. This is as follows:

$$4—1 \ldots 3—1 \ldots 2—1 \ldots 1—1 \ldots 1—2 \ldots 1—3 \ldots 1—4$$

(relative positivation)   *(relative negativation)*

In contrast to the suppressive action of the salts in which these same salt symbols appear in two separate series (see p. 349), they are here arranged in one continuous series.

One can make this difference clear as follows:

In the salt resistance we are dealing with the suppression of the interaction of colloid anions and colloid cations.

Although other factors also play a part, this interaction (at least at the optimum mixing proportion) will have the character of a product. The intensity of this interaction will thus only be able to decrease when salts are added. Because of the smallest shielding action of monovalent ions a salt of the type 1—1 (for example KCl) is thus always the least active. Increase of the valency of the cation or of the anion will cause the intensity of the interaction of the two colloid ions to decrease more strongly than 1—1 does. In consequence of this the double valency rule makes its appearance.

The electrophoretic charge on the coacervate surface has on the other hand the character of an algebraic sum since this surface consists of a mosaic of gelatin cations and arabinate anions. An added salt will, because its two ions have a shielding action on the oppositely charged colloid ions, necessarily bring about a change in the electrophoretic velocity. Whether an indifferent salt, for example of the type 1—1 (KCl), will bring about an absolute positivation or negativation cannot be stated *a priori* since this depends on whether the colloid anion is shielded more strongly by K· than the colloid cation by Cl' or *vice versa*.
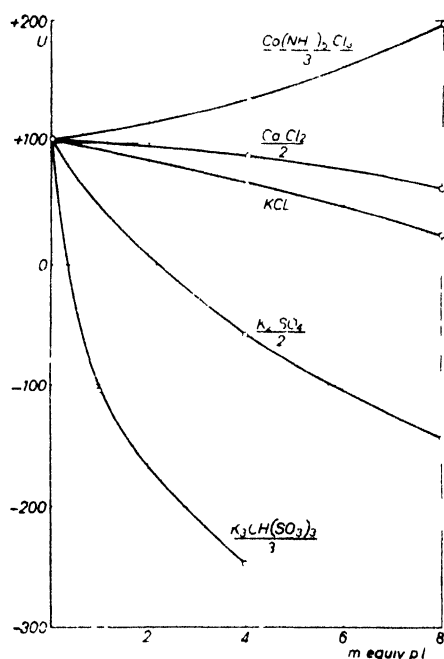
Fig 14. Continuous valency rule in the influence of added salts on the electrophoretic velocity of complex coacervate drops (gelatin and gum arabic)
Ordinates: electrophoretic velocity in arbitrarily chosen units
Without added salt the drops were charged weakly positive at the given mixing proportion and pH.
Abscissae: concentration of added salt in m. eq. per l.

If however one compares several salts with each other and at equal concentrations (in milli eq. per l.) it is to be expected that the coacervate drops will always come out relatively more positive with a salt of the type 2—1 than with 1—1, conversely always relatively more negative with a salt of the type 1—2 than with 1—1. If one extends this argument to still other terms on either side, one has the above given "continuous valency rule".

Under certain circumstances the continuous valency rule can also occur in experiments in which the suppressive action of neutral salts is investigated[1]. (See fig. 15A and C).

If namely the actual mixing proportion lies far removed from the optimum mixing proportion the coacervation is only weak. The first quantities of an added salt can now give rise to an improvement in the coacervation by preferential shielding of that colloid component which is present in excess.
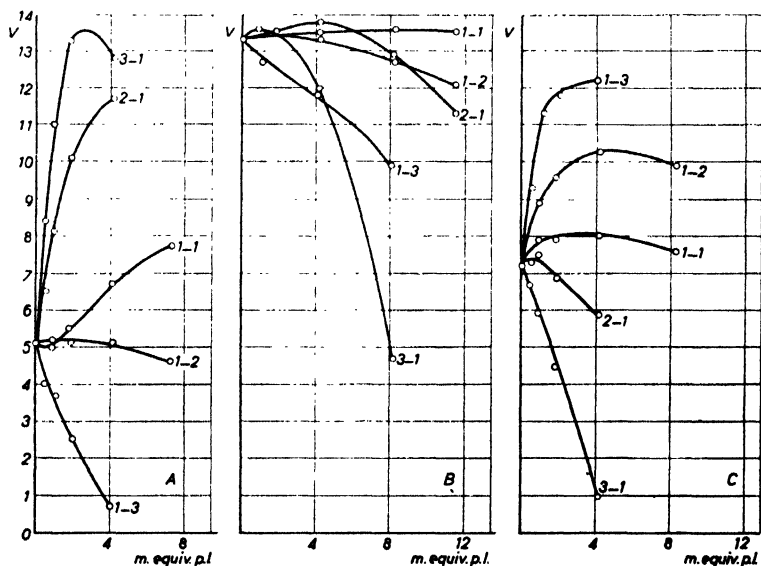
Fig. 15. Influence of small salt concentrations on the coacervate volume V of a strongly negatively charged (A), a practically uncharged (B) and a strongly positively charged (C) complex coacervate (see text).

If, for example, the coacervate is strongly positively charged (Fig. 15C) an excess of gelatin is present and now a small concentration of a salt of the type 1—3 causes improvement in the coacervation because the monovalent cation hardly shields the arabinate anion, the trivalent anion however shields the gelatin cation very much

    [1] H. G. BUNGENBERG DE JONG and C. V.D. MEER, Proc. Koninkl. Nederland. Akad. Wetenschap., Amsterdam, 45 (1942) 490.

more strongly. This leads to the originally unfavourable mixing proportion of the colloids now becoming better: the coacervation first increases; at higher salt concentrations however the suppressive action preponderates completely and the coacervation decreases to zero (not drawn). In principle the improvement will also exist with 1—2, although not to that extent. On the other hand salts of the type 2—1 and to a still greater extent the type 3—1 will act in the exactly opposite manner, because the originally unfavourable mixing proportion becomes still more unfavourable through the preferential shielding of the arabinate anion. The coacervation will with these salts therefore decrease directly even with the smallest concentrations. The result of it all is that the intensity of the coacervation at small salt concentrations is influenced according to the continuous valency rule.

Further, as indeed Fig. 15A shows, the sequence of the salt symbols in the bundle must be just the other way round if the coacervate is strongly negatively charged.

Fig. 15B gives the influence with an almost uncharged coacervate. The coacervate volume is here much greater than in A and C and here the pronounced fan-wise bunching of the curves, such as is present in A and C, is missing. In practice one can say that the salt influence merely consists of suppressive action. (The KCl curve, at somewhat greater concentrations than were investigated here, will also begin to drop, compare Fig. 12, p. 351).

Fig. 16. Continuous valency rule and double valency rule in the action of salts on a positively charged complex coacervate (gelatin + soya bean phosphatide). Ordinates: turbidity. Abscissae: salt concentration in m. eq. per l. The double valency rule makes its appearance in the actual suppressive action. Preceding this in small concentrations a bunching out according to the continuous valency rule (see text).
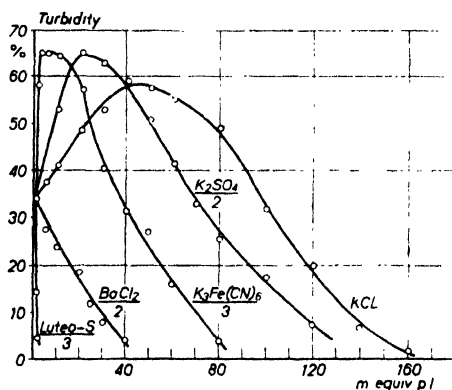
To avoid misunderstanding we must add that if on adding small concentrations of salt a bunching according to the continuous valency rule is first produced (such as for example Fig. 15A and C), nevertheless at higher concentrations, at which one is concerned with the real suppression, the curves are arranged according to the double valency rule.

See for example Fig. 16, which relates to a positively charged complex coacervate of gelatin and a soya bean phosphatide[1].

g. Composition of coacervate and equilibrium liquid in an isohydric series of mixtures (schematic).

For the purpose of a review of the analytical results of coacervates and equilibrium liquids one can with advantage make use of the diagrams customary in the Phase

[1] H. G. BUNGENBERG DE JONG and R. F. WESTERKAMP, Bioch. Z., 234 (1931) 367.

Theory. We consider then the mixtures of gelatin and gum arabic sols provisionally[1] as ternary systems, with gelatin (G), gum arabic (A) and water (W) as components.

If one plots the composition of coacervates and equilibrium liquids on an isotherm (compare p. 257, Fig. 12b in Chapter VIII, where an equilateral triangle WGA is used), it is seen that the region in which complex coacervation occurs lies very much to one side in the neighbourhood of the water corner[2].

In what follows therefore only a small portion of the phase triangle is depicted (from now on a right angled triangle is employed), in the neighbourhood of the corner W. The analyses which we discuss first relate to mixtures of isohydric[3] 2% gelatin and 2% gum arabic sols. The composition of these mixtures as totals, that is to say, ignoring whether coacervation takes place or not, thus lie on a line which connects the 2% G point on the ordinate axis to the 2% A point on the abscissa axis in Fig. 17 and Fig. 18.

Fig. 17. Scheme for the composition of coacervates (c_1, c_2 etc.) and equilibrium liquids (e_1, e_2 etc.) in an isohydric series of mixtures of 2% sols (from the mixtures m_1, m_2 etc.) (see text).

The closed curve obtained by connecting the branch of the equilibrium liquids with the branch of the coacervates, is not drawn to scale. Its real form is less symmetric and its position still depends on the pH (compare p. 359, Fig. 20).

Thus m_1, m_2, m_3, m_4 and m_5 represent the like mixtures. In the Phase Theory it is stated that when such a total system splits into two coexisting phases, the points which give the composition of these phases must lie together with that of the total system on a straight line.

The analytical results then show that complex coacervation may certainly not be considered as the separation of a gelatin-arabinate hydrate of constant composition (one could indeed imagine that this salt comprises such a weakly constituted lattice that it practically possesses the properties of a liquid). If this latter were correct the complex coacervate would have to have always the same composition (G content, A content, W content) independently of the chosen mixing proportion of the isohydric sols and thus must be depicted always by one and the same point in the triangle (for example point c in Fig. 18).

The equilibrium liquids e must then naturally vary in composition (because the points c, m and e must always lie on a straight line — see above).

The isotherm would thus have to exhibit approximately the shape of Fig. 18

---

[1] We shall see later that these mixtures in reality certainly cannot be considered as ternary systems. See p. 366, § 2m.

[2] In mixtures of gelatin and gum arabic sols still another coacervation region also occurs, situated more in the middle of the triangle. This coacervation is of a quite different kind than the complex coacervation and has already been discussed, see p. 255, Ch. VIII § 8.

[3] Isohydric liquids are liquids which possess the same pH. The A and G sols were brought to the same pH with HCl. The pH does not change on mixing the A and G sols. The pH of coacervate and the corresponding equilibrium liquid is also the same.

In reality (schematized in Fig. 17) one finds however that the complex coacervates, which are produced in an isohydric series of mixtures, do not have a constant composition but vary both as regards their G and A and their W content.

The isotherm therefore exhibits not only a branch on which the equilibrium liquids are situated but also a branch on which the coacervates (c_1, c_2 etc.) lie. Furthermore the coacervate branch bends on either side towards the corner W and the equilibrium liquid branch on either side away from the corner W and in addition from the G or the A axis. From this it is clear that the two branches connect up and the coacervation region is completely comparable with the closed regions of limited solubility as we know them in micromolecular ternary systems.

Although the analytical results give therefore no support for the conception of complex coacervation as simple formation of a gelatin-arabinate hydrate of constant composition, it will nevertheless be seen from the following subsections h - k that there are indications of a tendency in this direction.
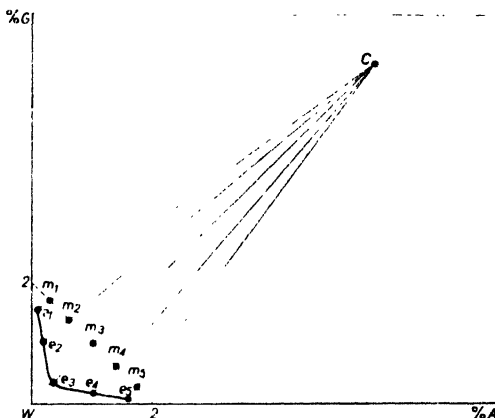
Fig. 18. Scheme for the composition of coacervate and equilibrium liquids in an isohydric series of mixtures of 2% sols, if the coacervate had a constant composition (see text).

h. Composition of coacervate and equilibrium liquid at the equivalent mixing proportion (schematic).

If in fig. 17 one produces the dotted lines, which connect coexisting coacervates and equilibrium liquids, downwards to the left, these lines in general intersect one of the coordinate axes.

Thus c_1, e_1 and c_2, e_2 intersect the G axis, c_4, e_4 and c_5, e_5 the A axis.

That is to say, in general the proportion of the two colloids in the coacervate and in the corresponding equilibrium liquid is not the same, since, if this proportion were indeed the same the produced dotted line through coacervate and equilibrium liquid would have to go always through the corner W. One sees from the figure that within an isohydric series of mixtures there is only one mixture of the two sols (in Fig. 17 that is m_3) in which the connecting line does go through the corner W (in this case therefore W, e_3, m_3 and c_3 lie on one straight line). This mixing proportion has particular significance for complex coacervation in view of the fact that it coincides very closely with two other extreme points in the isohydric series of mixtures, firstly with the mixing proportion at which the coacervate drops do not move electrophoretically, secondly with the mixing proportion at which the degree of coacervation (see below) is a maximum in the isohydric series of mixtures. We shall discuss both points briefly in more detail.

The reversal of charge phenomena in mixtures of two colloids have already been discussed (p. 321-327, Ch. IX, § 6a-c) to which reference may be made.

We saw there that at a given pH the reversal of charge point is met with at a definite mixing proportion of the colloids, independently of their absolute concentrations (compare p. 327, Fig. 46 and 47).
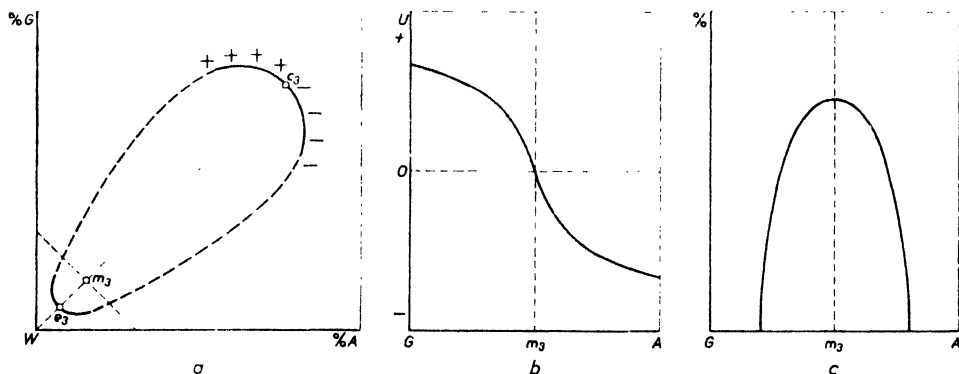
Fig. 19. Schemes for the discussion of the composition of coacervate and equilibrium liquid at the equivalent mixing proportion (see text).

a) + and —: signs of the charge of the coacervates, c_3 the equivalent coacervate (lies on a straight line with m_3, e_3 and the corner W).
b) Electrophoretic velocity as a function of the mixing proportion.
c) Degree of coacervation as a function of the mixing proportion.

The reversal of charge point is thus an equivalence point. Now this reversal of charge point practically coincides with the above mentioned mixing proportion m_3 (see scheme Fig. 19b) whereby we are justified from now on to denote this mixing proportion as the *equivalent mixing proportion* of the isohydric series of mixtures. See also the scheme Fig. 19a in which the sign of the charge is indicated along the coacervate branch.

We see that with mixing proportions relatively richer in the positive complex component (G) than the equivalent mixing proportion, the coacervate drops, are charged electrophoretically positive, with mixing proportions relatively richer in the negative complex component (A) are on the other hand charged negatively.

To characterise the maximum coacervation in an isohydric series of mixtures one can choose provisionally all kinds of criteria, for example, maximum turbidity, maximum volume of the coacervate layer, etc. It is however desirable to have at one's disposal a criterion which is calculable from the analytical results and in which the water content of the coacervate layer itself plays no part.

The attainment of an extreme value of the *degree of coacervation* can serve for this purpose. By this term we mean the fraction (expressed in %) of the two colloids (A + G) present in the total system which are to be found in the coacervate.

This quantity can be calculated from the A and G contents of the total system, of the coacervate and of the equilibrium liquid. The curve of the degree of coacerv-

ation in an isohydric series of mixtures has in general the shape as given in the scheme Fig. 19c.

With mixing proportions which are rich in G or rich in A there is no coacervation. The degree of coacervation is therefore zero there. In the range of mixing proportions where there is coacervation the degree of coacervation rises rapidly from zero on increase of the mixing proportion, expressed in % A, reaches a maximum at a certain mixing proportion and afterwards decreases again to zero. It now appears that the mixing proportion at which the degree of coacervation is a maximum almost coincides with the above discussed equivalent mixing proportion (m_3).

At this mixing proportion the positive gelatin ions therefore unite with the negative arabinate ions in the ratio of their equivalent weights corresponding to the given pH; here neither an excess of the one nor of the other colloid component is present. For this equivalent mixing proportion it is appropriate to speak of salt formation; a salt (gelatin-arabinate), however, which separates out with a fairly large amount of water not as a crystalline phase but as a typical liquid (in Fig. 17, p. 356, the coacervate c_3). The equilibrium liquid e_3 — in which the colloid ratio is the same as in the coacervate c_3 — then takes on the significance of the saturated solution of this colloid-colloid salt. In § 2 j (p. 360) we discuss further the composition of coacervate and equilibrium liquid at mixing proportions which deviate from the equivalent mixing proportion.

i. Position of the "demixing" region at various pH's

Fig. 20 reproduces the analytical results for some five isohydric series of mixtures of 2% gelatin and gum arabic sols. Of the 5 corresponding branches of the equilibrium liquids only two are drawn so as not to overload the figure. Here we already see that the "demixing" region is displaced on changing the pH.
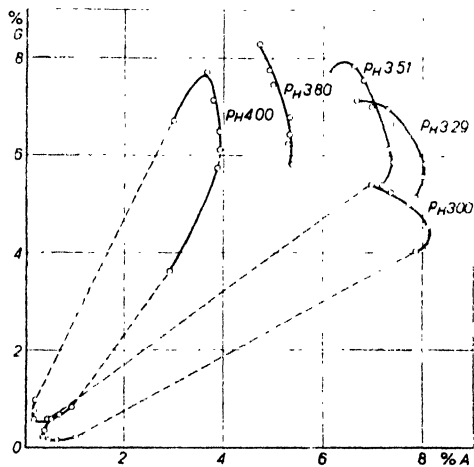
Fig. 20. Composition of coacervates and e-quilibrium liquids in some five isohydric series of mixtures of 2% gelatin and gum arabic sols. Top graph: coacervates lie on the 5 branches of the curves top right in the graph. Equilibrium liquids lie on 5 branches of curves near the water corner. In order not to overload the figure only two of these latter curves are drawn and further, apart from four, all tie lines, which connect coacervates to coexisting equilibrium liquids, are omitted. Bottom graph: The neighbourhood of the water corner drawn on an enlarged scale, containing the 5 branches of the curves of the equilibrium liquids.
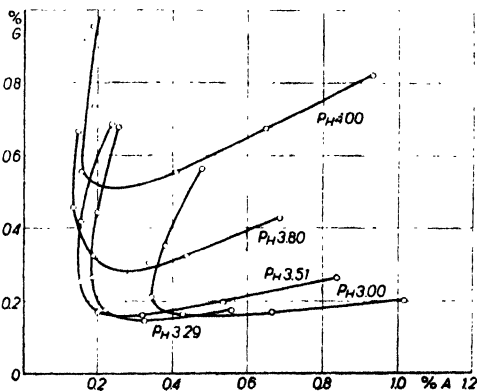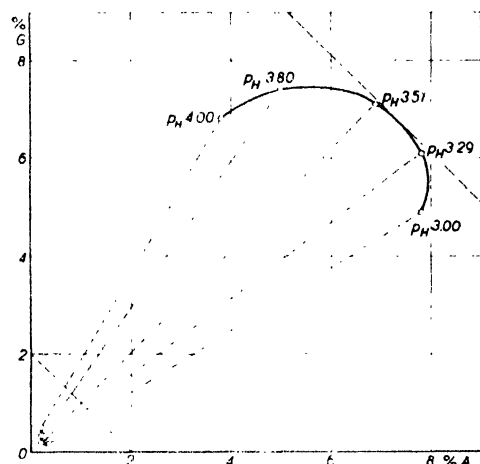
Fig. 21. The equivalent coacervates and equilibrium liquids corresponding to various values of pH (see text)

In Fig. 21 are plotted the compositions of the coacervates and equilibrium liquids which represent the equivalent mixing proportion in each of the 5 isohydric series of mixtures.

The dotted lines are thus analogous to $e_3 - m_3 - c_3$ in the scheme Fig. 17 (p. 356).

We see that the slope of the dotted lines becomes smaller on lowering the pH, which means that the equivalent mixing proportion is displaced towards higher percentages of A by lowering the pH.

This is also quite to be expected because the equivalent weight of the gelatin increases on lowering the pH and that of the gum arabic decreases (see p. 323).

Compare the displacement of the reversal of charge point with the pH in Fig. 6 (p. 346) and in Ch. IX, Fig. 43 (p. 324).

j. Distribution of the colloid component present in excess in an isohydric series of mixtures over coacervate and equilibrium liquid

This distribution is less easy to survey from graphs such as those used in Fig. 17 and Fig. 20 since in these graphs three quantities W, G and A vary simultaneously.

It is more easy to survey the various points when we set aside for the moment the simultaneously varying water content and thus restrict ourselves to the colloid compositions of the total mixture, the coacervates and the equilibrium liquids[1].

The following table gives the analytical results for an isohydric series of mixtures at pH 3·51. Here one finds from left to right successively the mixing proportion of 2% sols, expressed in % of gum arabic sol (100 A/(A + G)), the gelatin contents

Mixing proportion of the sols	Composition of coacervate		Composition of equilibrium liquid		100 A/(A + G)	
100 A/(A + G)	% G	% A	% G	% A	in coacervate	in equilibr. liquid
33.6	7.69	6.11	0.686	0.234	44.3	25.4
40.4	7.86	6.60	0.419	0.151	45.6	26.3
45.4	7.55	6.77	0.256	0.154	47.3	37.6
50.4	7.01	6.93	0.170	0.200	49.7	54.1
55.4	6.15	7.30	0.162	0.318	54.3	66.3
60.4	5.88	7.38	0.195	0.535	55.7	73.3
65.3	5.41	7.15	0.261	0.839	56.9	76.3

[1] H. G. Bungenberg de Jong and W. A. L. Dekker, Kolloid Beih. 43 (1936) 213, see p. 233 and fig. 11.

H. G. Bungenberg de Jong, Proc. Koninkl. Nederland. Akad. Wetenschap., Amsterdam, 50 (1947) 707.

(% G) and gum arabic contents (% A) of coacervate and equilibrium liquid. In the two last columns of the table one now finds the colloid compositions of the coacervate and the equilibrium liquid calculated from the above, both expressed in % A ($=$ 100 A/A + G).

In Fig. 22 the colloid compositions of coacervate (column 6) and equilibrium liquid (column 7) are plotted as a function of the mixing proportion of the two sols (column 1). Two separate curves result, C for the coacervate, E for the equilibrium liquid. These curves intersect at the equivalent mixing proportion. Here or at any rate nearly here lies the electrophoretic reversal of charge point of the coacervate.

In the figure three auxiliary lines have also been added. The one goes vertically through the equivalent mixing proportion. This divides the plane of the figure into a left half, in which the coacervate surface is electrophoretically charged positive and into a right half in which the coacervate surface is electrophoretically negative.

The second auxiliary line ("Equiv."), drawn horizontally through the equivalence point, reproduces the colloid composition of the equivalent coacervate.

Since the mixtures were prepared by mixing two equally concentrated sols, the colloid compositions of the total mixtures lie on the third added auxiliary line ("M") which goes through the origin at an angle of 45°.

Since at the equivalent mixing proportion the colloid composition of the coacervate is the same as that of the equilibrium liquid, it is thus also equal to that of the total mixture. The third auxiliary line drawn at an angle of 45° thus also goes through the intersection of the curves C and E.

Fig. 22. Changes in the colloid composition of complex coacervate and equilibrium liquid in an isohydric (pH 3.5) series of mixtures of equally concentrated (2%) sols of gelatin (G) and gum arabic (A). Ordinates: the colloid compositions expressed in % A in the ratio of A to A + G (thus 100 A/(A + G)). Abscissae: mixing proportion of the isohydric sols expressed in % of the A sol (100 A/(A + G)). The colloid compositions of the total mixtures lie on the dotted straight line M, drawn at an angle of 45° towards the top right corner. Curves C and E give the colloid compositions of the coacervates and equilibrium liquids. At the equivalent mixing proportion the curves C and E intersect, the colloid composition of both is necessarily equal to that in the total mixture. At other mixing proportions of the sols the colloid compositions of coacervate, equilibrium liquid and total mixture are no longer equal. On comparison of the position of the C and E curves with respect to the horizontal dotted line (which is the colloid composition of coacervate and equilibrium liquid at the equivalent mixing proportion) it appears that at excess of one of the colloid components the colloid composition of the coacervate changes least, while that of the equilibrium liquid changes most. See further text.
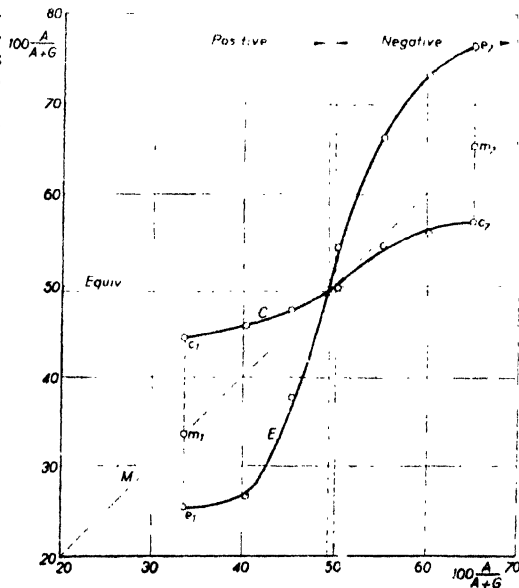
From the situation of these C and E curves with respect to the horizontal auxiliary line it appears that, with an excess of one of the colloid components, the latter always distributes itself over the two liquid layers in favour of the equilibrium liquid. Thus the mixture m_7 splits into an equilibrium liquid e_7 which is relatively still richer in A than m_7 and into a coacervate c_7 which is relatively little richer in A than the equivalent coacervate. The same is to be seen with positive coacervates: the equilibrium liquid e_1 is still richer in G than m_1, but the coacervate c_1 is but little richer in G than the equivalent coacervate.

From this much smaller displacement of the coacervate composition than that of the equilibrium liquid there is seen an attempt by the processes active in the coacervation to maintain the separation of the equivalent coacervate. In consequence of this the colloid component present in excess goes mainly into the equilibrium liquid. Nevertheless a smaller fraction of this component does penetrate into the coacervate and thereby brings about a change of the composition whereby on the one hand the uncharged equivalent coacervate assumes at its surface the electrophoretic sign of the charge of the colloid component present in excess (see p. 358, Fig. 19b), on the other hand the mutual solubility of coacervate and equilibrium liquid increases (compare the decrease of the degree of coacervation in Fig. 19c, p. 358).

Fig. 22 holds for some definite pH. If the pH of the isohydric series of mixtures is chosen differently, the system of intersecting C and E curves is so displaced that their intersection always remains on the dotted oblique line through the origin. At higher pH this intersection is displaced downwards towards the left, at lower pH upwards towards the right. For the rest the statements set out above hold just the same for the relative situation of the curves and the conclusion drawn from it.

k.   Changes of composition of coacervate and equilibrium liquid
at constant mixing proportion and variation of the pH

By combining the analytical results of the 5 isohydric series of mixtures investigated it is possible to survey how the G and A and W content changes at constant mixing proportion with variation of the pH[1]. We choose 50% as the mixing proportion since only this mixing proportion occurs in all 5 series of mixtures and then find the colloid contents stated in columns 2, 3, 4 and 5 of the following table.

From these were calculated the colloid ratios in coacervate and equilibrium liquid (column 6 and 7).

| pH | Coacervate | | Equil. liquid | | 100A/(A + G) | | Mixing proportion of equivalent coacervation |
	% G	% A	% G	% A	Coacervate	equil. liquid	
3·00	5.44	6.92	0.59	0.495	56.0	45.6	62.2
3·29	6.68	7.39	0.277	0.183	52.5	39.8	55.2
3·51	7.06	6.92	0.175	0.195	49.5	52.7	49.5
3·80	6.25	5.27	0.324	0.446	45.7	57.8	40.4
4·00	3.9	3.0	0.80	0.90	43.5	52.9	32.7

[1] H. G. BUNGENBERG DE JONG and B. KOK, Proc. Koninkl. Nederland. Akad. Wetenschap., Amsterdam, 45 (1942) 51.

Fig. 23. Composition of coacervates and equilibrium liquids at constant mixing proportion (1 : 1) of the 2% gelatin and gum arabic sols, but with the pH varying (see text).

The dotted line touching the branch of the coacervates at an angle of 45° gives compositions which are equally rich in water. Any line parallel to this which lies closer to the corner gives compositions which again each have the same water content but in which the water content is greater than that of the points on the dotted line. The coacervates at pH 3.8 and 3.0 are thus richer in water than those at pH 3.3 and 3.5. The coacervate at pH 4.0 is still more rich in water.

If one plots the data of columns 2, 3, 4 and 5 in a graph fig. 23 is produced, from which it follows that W, A and G all change when the pH is varied at constant mixing proportions.

As in § 2 j. we shall here also provisionally ignore the change in the W content and thus restrict ourselves to the colloid ratios in coacervate, equilibrium liquid and total mixture[1] (see Fig. 24).
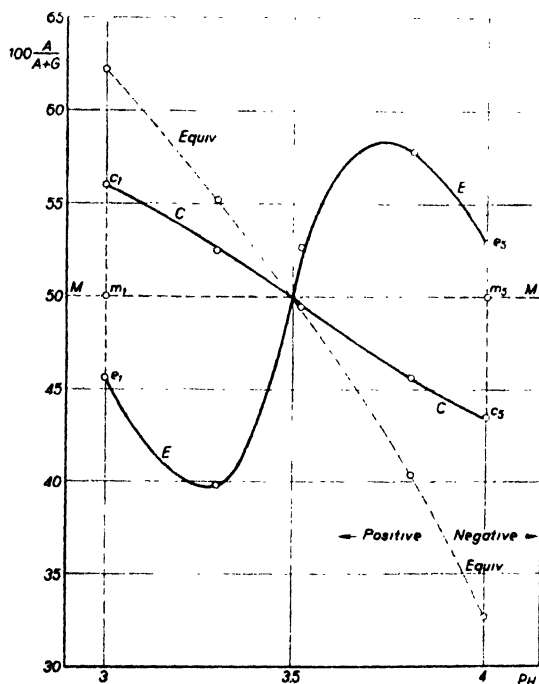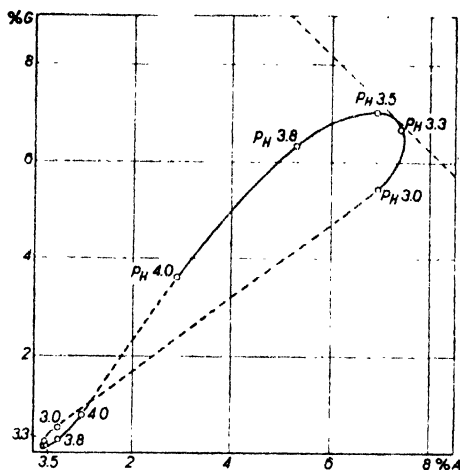
Fig. 24. Changes in the colloid composition of complex coacervate and equilibrium liquid on varying the pH while the mixing proportion (1 : 1) of the 2% gelatin and gum arabic sols remains constant.
Ordinates: as in Fig. 22 (see p. 361).
The colloid compositions of the total mixtures lie on the horizontal dotted line, those of the coacervates on curve C, those of the equilibrium liquids on curve E.
At the pH, at which the mixing proportion 50% is the equivalent, the C and E curves intersect and this intersection lies necessarily on the horizontal dotted line.
At other pH values these three colloid compositions are no longer equal.
The dotted curve rising obliquely to the left gives the colloid compositions of the equivalent coacervates (and equilibrium liquids) corresponding to the various pH values.
On comparison of the position of the C and E curves with respect to that dotted curve it appears that here we have a behaviour of the complex coacervation similar to that which appeared in Fig. 22 that is to say, there is a tendency for the separation of the equivalent coacervate corresponding to the given pH. See further text.

[1] H. G. BUNGENBERG DE JONG, *Proc. Koninkl. Nederland. Akad. Wetenschap.*, Amsterdam, 50 (1947) 707.

These colloid compositions (columns 6 and 7) are reproduced as functions of the pH by the curves C and E. The colloid proportions of the total mixtures lie on a horizontal line M (dotted) because they only refer to the same mixing proportion (50% A).

In Fig. 24 also we encounter two intersecting curves C and E, the intersecting of which lies on the dotted horizontal line. This intersection lies at that pH at which the mixing proportion of the sols is just the equivalent one. The colloid proportion in coacervate, equilibrium liquid and total mixture is here the same. At other pH's these three are all different.

In the figure is also drawn a curve ("Equiv.") on which the colloid compositions of the equivalent coacervates, corresponding to various pH's, lie. (column 8). Since here or very near by lies the reversal of charge, that curve divides the plane of the figure into a positive (left) and a negative half (right).

From the relative situation of this dotted curve ("Equiv.") running obliquely towards the left and the two other curves drawn (C and E), one sees that a behaviour of the complex coacervation is present here which is similar to that discussed above in § 2 j: there is always a tendency for the separation of the equivalent coacervate corresponding to the chosen pH.

At constant mixing proportion this tendency can only be satisfied at one pH value. At the remaining pH values this is not the case but this tendency is manifested by the fact, that from a given total mixture (for example m_1 or m_5) a coacervate (c_1 or c_5) is formed which lies considerably closer in colloid composition to the equivalent coacervate corresponding to that pH than the total mixture; and an equilibrium liquid (e_1 or e_5) which lies appreciably further away from it.

l. Influence of an added salt on the composition of coacervate and equilibrium liquid

In § 2f, p. 349 we already saw from qualitative experiments that salts in general exert two kinds of action: on the one hand a suppressive action in which the so-called "double valency rule" holds and a displacement of the optimum mixing proportion in which the so-called "continuous valency rule" holds.

According to analytical results of isohydric series of mixtures with and without added salt, the two influences make themselves felt in the composition of coacervate and equilibrium liquid. See Fig. 25, which holds for a series of mixtures of 2% sols. The influence of the added salt manifests itself first of all in a mutual approach of the coacervate branch and the equilibrium liquid branch (of the same closed "demixing" curve, p. 357, § 2 g). The former is displaced in the direction of the corner W, that is to say, the coacervates become richer in water. The latter branch is displaced away from the corner W, that is to say, the equilibrium liquids increase in colloid content. *The mutual solubility of coacervate and equilibrium liquid is therefore increased by the added salt.* At still higher $CaCl_2$ concentrations the closed coacervation region will contract further (and will finally disappear in a point) and as soon as it is no longer cut by the line connecting 2% G and 2% A on the coordinate axes of the phase diagram (compare p. 356, Fig. 17), the added salt will have "suppressed" the coacervation in the isohydric series of mixtures of 2% sols.

Besides this displacement of the curves in the direction of the corner W (coacervates) or away from it (equilibrium liquids) still another systematic displacement is present. This manifests itself in a decrease in the slope of the line which connects the corner W to that coacervate (indicated by an arrow) which corresponds in each series of mixtures with the maximum degree of coacervation (p. 358, § 2 h). Calculation from the analytical figures gave for these optimal mixing proportions: blank = 48% A; 5 and 7.5 m. equiv. p. 1 = 55 and 56% A.

Fig. 25. Influence of CaCl$_2$ on the composition of coacervate and equilibrium liquid produced in isohydric series of mixtures (pH 3.5) of 2% gelatin and arabinic acid sols. a: the arrows on the coacervate curves give the position of the coacervate corresponding in each series of mixtures with the maximum degree of coacervation.
b: the neighbourhood of the water corner with the curves of the equilibrium liquids on a larger scale.
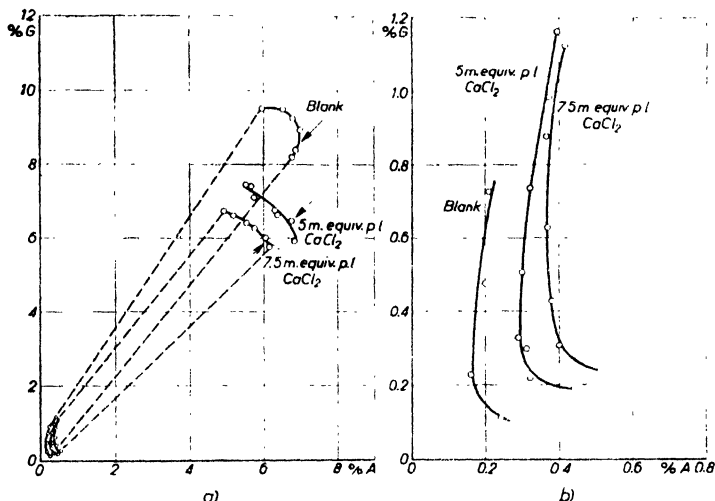
For this series the usual gum arabic sol (adjusted with HCl to pH 3.5) had been replaced by arabinic acid sol (adjusted with NaOH to pH 3.5), thus not containing Ca ions. This displacement did in principle not alter the general characteristics of the usual complex coacervation, but it hindered greatly the formation of two clear macrolayers (necessary for sampling) from coacervated systems with a A content greater than corresponds with the maximum degree of coacervation. Therefore the figure shows mainly the composition of positively charged coacervates and corresponding equilibrium liquids.

As the reversal of charge points lie close by the arrows one sees here a positivating influence arising from the added CaCl$_2$. Indeed at a mixing proportion of 51% A the coacervate is charged weakly negative in the blank series, on the other hand weakly positive with 5 and 7.5 m. eq. per 1. CaCl$_2$.

Thus in view of the "continuous valency series" (p. 353, § 2 f) previously discussed one can expect a similar but still stronger displacement for an added salt of the type 3—1, on the other hand displacement to the other side for a salt of the type 1—2 and still stronger for a salt of the type 1—3, while a salt 1—1 will have practically no influence on the position of the reversal of charge and of the maximum degree of coacervation in the isohydric series of mixtures. Practically only the first of the influences discussed here will exist in the case of a salt of the type 1—1, that is to say, the suppressive action.

m. Influence of the colloid concentration of the isohydric sols

As yet only the results concerning series of mixtures of 2% sols have been stated. In the scheme Fig. 26 the curves C and E may once more belong to an isohydric series of mixtures of 2% sols (the curves represent the two readily accessible portions of one and the same closed "demixing" figure, p. 356, § 2g).
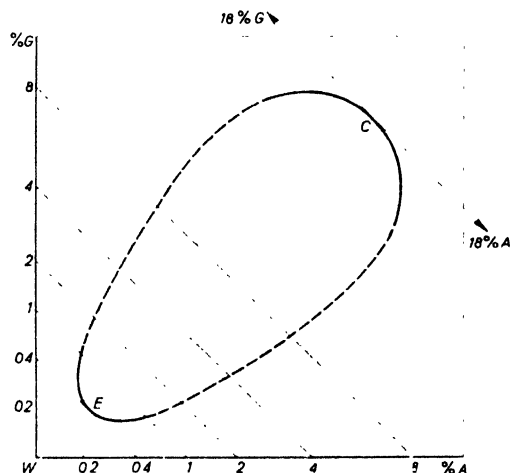
Fig. 26. Scheme in the discussion of the influence of the colloid concentration of the isohydric sols on complex coacervation[1]). See text.

Tangents have been drawn to these two curves at an angle of 45°. These intersect the coordinate axes at about 18% (tangent to the C branch) and at about 0.4% (tangent to the E branch).

Now one must expect that on choosing another concentration of the two isohydric sols:

1st no coacervation will occur when the colloid concentrations are lower than 0.4% (for example 0.2% sols see in the figure);

2nd coacervation always occurs in a certain range of mixing proportions when the concentration of the isohydric sols lies between 0.4% and 18% (for example 1, 4 and 8% sols: see in the figure).

Now experience already shows clearly that neither of the two expectations comes true.

On the one side 0.2% sols do give genuine coacervation; one can even choose the colloid concentrations lower and lower still without coacervation ceasing, for example with 0.01% sols. A lowest limit has not been found here and must lie lower than 0.001% sols (the practical limit of observation lies in this region).

As regards the second expectation coacervation ceases already long before 18% on increasing the sol concentrations.

With 5% and more highly concentrated sols coacervation no longer occurs.

From these simple facts one already arrives at the conclusion that there is something which does not fit with the simple idea that complex coacervation should be comparable with an "demixing" in a ternary system W—G—A.

The nature of the deviations is shown by analyses of coacervates and equilibrium liquids of isohydric series of mixtures with various colloid percentages of the sols. See Fig. 27 which reproduces the results for 1, 2, 4, 4.5 and 5% series of mixtures (the same pH throughout).

If now we really had to do with a ternary system, then, independently of the chosen colloid concentration, only one branch of a curve on which the coacervates lie and one branch of a curve on which the equilibrium liquids lie ought to be found.

---

[1] The schematic figure has purposely not been drawn in the correct proportions.

The results show that there is no question of this. With each colloid concentration a different position of the C and E branch of the "demixing" figure is found. One can now ask whether the Phase Rule does not hold in principle for our colloid systems. This question had indeed significance as long as sols of colloids such as gelatin and gum arabic were still considered as two-phase systems.

With the modern conception of these sols as true solutions of high molecular substances this objection falls and thus there only remains the consideration whether we have correctly enumerated the number of components (see also small print, p. 253, Ch. VIII, § 6).

In the preceding sections we have already arrived at the interpretation that in complex coacervation colloid-colloid salt is formed. The essential partners in the complex coacervate are then gelatin cations and arabinate anions.

Yet the gelatin or gum arabic sols, from the mixing of which the coacervation is produced, contain not only the ions mentioned but also an equivalent amount of ions of the opposite sign.

Fig. 27.   Influence of the colloid concentration of the isohydric sols (pH 3.5) on the composition of coacervates and equilibrium liquids. See text.

The gelatin sol is a solution of gelatin chloride, the gum arabic sol a solution of Ca arabinate [1].

Phase Theory shows however that a system of the type: water — salt AB — salt CD, that is to say, in which the two salts possess no common ion, is not a ternary but a quarternary system. Thus at constant temperature and pressure it possesses one more degree of freedom than a ternary system. This is therefore applicable to complex coacervation and it should not surprise us that in the "only" ternary diagram of Fig. 27 a displacement of the C and E curves occurs on increase of the colloid concentrations.

Let us now consider more closely in what *direction* these branches of the curves are displaced. One sees from the figure that according as the colloid concentration of the series of mixtures lies higher the C and E branches approach more closely to one another.

The "demixing" region therefore contracts on increase of the colloid concentration As a result of this the finding becomes understandable that even in a 5% series of mixtures coacervation no longer occurs. Conversely on lowering the colloid concentration below 1% the "demixing" region must expand, that is to say, the C branch then lies still further away from the corner W and the E branch lies still closer to the corner W. Through this it becomes understandable that in series of mixtures of 0.4% and lower coacervation can still always occur.

---

[1] Abstraction made of a minor content of other cations; e.g. Mg, K, Na.

Now it is striking that the displacements discussed above are of the same nature as we saw in § 2 1 (p. 364) for the influence of an added indifferent salt.

Similarly on increase of the colloid concentration there occurs a displacement of the mixing proportion at which the maximum degree of coacervation and the reversal of charge point lie, in the same direction as that which we saw for constant colloid concentration and increase of the $CaCl_2$ concentration (see p. 365, Fig. 25).

In brief the effects of an increase of the colloid concentration in an isohydric series are the same as those of added $CaCl_2$.

This is however not at all surprising, since this really follows already from the interpretation which we have given of complex coacervation as the formation or the tendency to the formation of an equivalent colloid-colloid salt.

The gelatin sol is a solution of gelatin chloride, the gum arabic sol a solution of (mainly) Ca arabinate. If the colloid cations and colloid anions unite with each other in the equivalent mixing proportion, then mutually equivalent amounts of Ca and Cl ions remain over.

Now an indifferent salt has a suppressive action on complex coacervation (p. 349, § 2f). It increases the mutual solubility of coacervate and equilibrium liquid. The concentration of the $CaCl_2$ produced in this double decomposition will be greater the higher the colloid concentrations of the two isohydric sols.

That the displacements, which occur in Fig. 27 as a result of an increase of the colloid concentrations of the sols, are of the same nature as the displacements in Fig. 25 as the result of addition of $CaCl_2$ is thereby understandable.

One can also calculate that these parasitic $CaCl_2$ concentrations are indeed of the same order of magnitude in the mixing of 5% sols as those at which a complex coacervate produced from very dilute sols is suppressed.

Also the fact mentioned in § 2f (p. 352) of the decrease of the indifferent salt resistance on increase of the colloid concentrations of the two isohydric sols becomes directly understandable. More $CaCl_2$ is in fact then formed in the double decomposition and one then needs add successively less $CaCl_2$ (or another salt) just to suppress coacervation.

n. The water content of the complex coacervates

With regard to the water content one can in general say that it is large with unfavourable conditions for complex coacervation and decreases in proportion as these become more favourable.

We can read this off from certain of the previous diagrams when one bears in mind that in a right angled triangular diagram in which the corner W is the right angle, lines sloping down at an angle of 45° from upper left to lower right, represent lines of constant water content.

When thus we draw in such a diagram a tangent at this angle to the coacervate branch the point of contact indicates that coacervate which has the minimum water content.

Let us first examine Fig. 23 on p. 363 in which such a tangent has been drawn. It touches the coacervate branch at or near that pH at which the line connecting coacervate and corresponding equilibrium liquid and produced downwards passes through the corner W, that is to say, at that pH the given mixing proportion of the colloids is also the equivalent one (the reversal of charge point also lies close by).

Thus the above stated fact is exhibited clearly in this figure; the water content is a minimum at the most favourable condition for complex coacervation (here a definite pH at the given mixing proportion). On decrease or increase of the pH the water content increases.

In a discussion of the indifferent salt resistances it also became manifest that in every isohydric series this is a maximum near the equivalent mixing proportion (near the reversal of charge point) but that this maximum resistance still depends on the pH and at about pH 3.5—3.7 reaches its highest value (see p. 352, Fig. 13).

Let us now compare Fig. 21 (p. 360) in which a tangent (dotted) has again been drawn to the coacervate branch. This coacervate branch gives the composition of the coacervates which are throughout the equivalent coacervates at the pH values indicated.

The tangent touches this coacervate branch at about pH 3.4—3.5. Of all the equivalent coacervates this one therefore contains the least water. The agreement with the pH's which follow from the maximum salt resistances is still fairly reasonable, the more so when one bears in mind that these latter were determined on sols of much smaller concentration.

Fig. 28. Dry weights of the complex coacervates as a function of the mixing proportion of the isohydric 2% sols. Ordinates: Dry weights in %, that is to say, gelatin + gum arabic content. Abscissae: Mixing proportion on % of the gum arabic sol. Arrows give the mixing proportions of optimum coacervation.

One would now expect that the water content of the coacervate is a minimum in every isohydric series of mixtures at the equivalent mixing proportion (reversal of charge point). In Fig. 28 the dry weights, that is to say, the A + G contents, of the coacervates for the various isohydric series of mixtures of Fig. 20 (p. 359) are plotted as a function of the mixing proportion. It is seen from these figures that the expectation is not in general borne out.

One would indeed expect for the shape of the A + G curves in the figure curves which possess steeply rising branches on both sides and have a top (maximum dry weight, that is to say, minimum water content) at the equivalence point (arrows in the figure).

It is true these curves possess steeply rising branches at the extremes (see for example that for pH 4.0) but the symmetrical top is replaced by a sloping plateau in which the equivalence point lies.

The minimum water content in the isohydric series of mixtures thus lies to the left of the equivalent mixing proportion, i.e. at mixtures which are relatively poorer in gum arabic.

It is possible that the explanation of this discrepancy lies in the fact that the accompanying ions of the gelatin, the Cl ions, are only monovalent, those of the arabinate ion, the Ca ions, however are divalent. We have already argued that in principle micro ions counteract the coacervation, that is to say, increase the water content of the coacervate (p. 364, § 2 1).

In an isohydric series with rising gum arabic content the concentration of the

gelatin decreases and therefore that of the Cl ions also, at the same time however the concentration of the gum arabic and thus of the Ca ions also increases. Passing through a series of mixtures from left to right a monovalent suppressive ion is then diminished and a divalent, and therefore more strongly suppressive, ion is augmented (influence of valency see p. 349, § 2 f). As a nett result the suppressive action must therefore increase from left to right, that is to say, the A + G content must decrease.

o. Theory regarding the internal state of the complex coacervate

All the foregoing facts indicate that the opposite charge of the gelatin (positive) and the gum arabic (negative) is the essential factor which unites the two colloids in the coacervate. Further one is struck by a fairly clear inverse correlation with the water content of the coacervate. All the factors which render the mutual electric attraction less effective (unfavourable mixing proportion, added indifferent salts, unfavourable pH) increase the water content. In the original theory of complex coacervation the water content of a coacervate was still conceived as water of hydration (p. 243, Ch. VIII, § 3) and we thus arrived at the pronouncement that the internal state of a complex coacervate is regulated by two opposing factors: effective attraction and hydration. The former attempts to make the kinetic units of both colloids approach each other as closely as possible, the latter opposes this in view of the fact that it amounts to a dehydration.

All later experience is in agreement with the essential importance of the first mentioned factor. Attempts however to find correlations which would connect up with a variable degree of "hydration" have had a negative result [1].

This led to a more cautious formulation of the two oppositely connected factors: "effective attraction of an electrostatic nature versus repulsion connected with the tendency to hydration".

The new insight into the nature of the sols of "hydrophilic" colloids as true solutions and into the structure of the kinetic units in the case that the colloids belong to the high viscous type (to which gelatin and gum arabic both belong) have changed our conceptions regarding the water content of coacervates (p. 248, Ch. VIII, § 4). This new insight makes it possible to perceive that

a. the water content of the complex coacervates is connected inversely with the magnitude of the effective attraction,

b. no hydrational repulsive factor in the true sense need be assumed,

c. nevertheless a is connected with hydration.

We therefore base our discussion on the idea that the kinetic units of the gelatin sol as well as of the gum arabic sol consist of statistically kinked macromolecules in true solution [2].

For this form to be assumed however it is necessary for a sufficient number of hydrophilic groups to be distributed in the length direction along the macromolecule (for example OH groups) which by hydration ensure that the macromolecule is bathed in solvent to a sufficient extent over its entire length, that is to say, is in solution over its entire length. To this extent what follows is connected with hydration

---

[1] H. G. Bungenberg de Jong, P. van der Linde and A. de Haan, Rec. trav. chim., 54 (1935) 17.

[2] To what extent this is really fully satisfied will not be considered here. See note p. 248 and small print on p. 257.

(point c). This true hydration is however to be sharply distinguished from the water which is to be found in the meshes of the kinked macromolecule. This water is not bound to the macromolecule and is therefore only *water of occlusion*. In gelatin and gum arabic the amount of occlusion water is very large compared with the true hydration water.

We must now further consider that there are also ionised groups along the macromolecule (positive with gelatin, negative with gum arabic).

As a result of this the "length of a chain element" (see p. 95, Ch. IV) will be larger than when they are not present (see Chapter VII, § 6, p. 210-211). The macromolecular clews are consequently less dense and the amount of occlusion water per clew is larger than in the absence of ionised groups. Now in the complex coacervate gelatin cations and arabinate anions are present in intimate intermixture. We also assume in this that both colloids are present in the clewed form but it may be foreseen that the clews will partly penetrate each other. The positively ionised groups of the gelatin will attempt to attach themselves to — or at any rate to approach very closely to — the negatively ionised groups of the gum arabic for which purpose loops of the one macromolecule will possibly penetrate between loops of the other. Let us now consider the state at the equivalent mixing proportion. Then inside the enclosed volume of a clewed molecule of gelatin as many positive as negative ionised groups will be present (the latter attached to loops of the arabinate) and these each time in pairs at a short distance from one another. The consequence of this for the gelatin molecule must be that a considerable reduction of the rarity of the clewed macromolecule will have occurred.

The state of expansion in which the positively charged gelatin clews existed in the original sol with respect to a fictitious gelatin without charges is lost after coacervation in view of the fact that now beside each positive charge there is a negative charge at a short distance away. It is even not inconceivable that a state of contraction with regard to this fictitious gelatin will set in. In fact the pairs of positive and negative charges present inside the clew will behave as electric dipoles and also exert a certain attraction upon each other.

So far we have considered the enclosed volume of a clewed molecule of gelatin in the coacervate. The same argument can be applied to the enclosed volume of a clewed molecule of the arabinate and one can here also arrive at the conclusion that this clew must be much less rarified than in the original arabinate in the sol state.

For the above argument we started from a coacervate at the equivalent mixing proportion. If one of the colloids is present in the coacervate in excess then all the positive and negative ionised groups of the macromolecules will no longer be able to unite into pairs but a certain amount of one sign will remain over (even though also accompanied by a micro ion of the opposite sign). It is clear that as a result of this excess the average rarity of the clews entangled in each other will be greater than at the equivalent mixing proportion. That means therefore that excess of one of the colloid components will cause the water content of the coacervate to increase, or as we expressed it in point (a): the water content of the complex coacervates is reciprocally connected with the magnitude of the effective attraction. Point (b) is seen from the above arguments: no hydrational repulsion factor need be assumed, indeed the true hydration has nothing to do with the change in the water content. These changes are nothing but changes in the amount of water of occlusion as a result

of changes in the rarity of the macromolecular clews. And these latter depend on a change of the "length of a chain element" as result of electrical causes.

The existence of an optimum pH for complex coacervation (p. 360, Fig. 21) appears to speak in favour of a fairly considerable interpenetration of the clews of both kinds. The equivalent coacervate at this pH is poorer in water than equivalent coacervates at other pH values. With strong penetration the pairs of positive and negative charges are to be found more or less uniformly throughout the enclosed volume of a macromolecule and the mutual attraction of these dipoles (which leads to a state of contraction) depends then also on their number. This number would then be the greatest at the optimum pH so that here the water content becomes a minimum.

The interpenetration of the clews need also not be in conflict with the ready suppressibility of coacervates by salts since this penetration only happens with the loops of the macromolecule and an unrestorable entanglement of them need not also be the consequence. Indeed the disintegration phenomena (p. 347-349) out of which the relative displaceability of gelatin and gum arabic appears, points also in that direction. In addition we have to assume that as a result of the heat motion these loops continuously change place and shape. So it becomes understandable that the complex coacervate, although fairly viscous, nevertheless behaves as a Newtonian liquid.

Finally we must also consider the increase of the water content of a complex coacervate by indifferent salts. This is based on a shielding of the negative ionised groups by the cation of the added salt and of the positive ionised groups by the anion. Through this the attraction between the ionised groups of the two colloids becomes smaller than it was originally, so that the macromolecules can indulge more in their tendency to distribute themselves over a large volume (increase of entropy), in other words the water content of the coacervate will increase.

In the above given considerations an attempt has been made from modern points of view to explain the changes in the internal state of the coacervate for the case where both colloids are of the statistical clewed type. They will also be applicable though changed more or less if only one of the colloids is of this type, the other of the globular type (for example, serum albumin (positive) + gum arabic (negative), p. 233, Fig. 2). This will still be the case even for the variants to be discussed in the next section, colloid cation + micro anion or colloid anion + micro cation (§ 3, p. 384). The existence of the variant micro cation + micro anion (p. 407, § 4) however raises new problems. We shall have to decide from the facts that considerations based on macromolecular structure, although they can be very useful, do not yet elucidate the essential point of the complex coacervation. See further p. 412, § 4 c.

p. Complex flocculation

We have already discussed (p. 233, Ch. VIII, § 1b) that flocculation merely represents the separation of a new phase in a highly disperse form, and that in many cases this phase is a coacervate. This is just as well applicable here. We know various cases in which flocculation occurs on mixing two oppositely charged sols, which has all the characteristics of complex coacervation: reversibility, rules regarding the influence of the mixing proportion and of the pH, reversal of charge phenomena, suppression by indifferent salts according to the double valency rule and occurrence of the continuous valency rule. In these cases we can speak of *complex flocculation*, whereby, although the manifestation of the liquid nature is missing, we nevertheless

have to reckon with the presence of the same sort of complex relations as we have discussed above in the case of complex coacervation.

In some cases it is also possible to demonstrate the coacervate nature of the complex flocculi. Thus in the combination positive gelatin + negative phosphatide[1] thoroughly liquid coacervate drops occur close by the isoelectric point of the gelatin or close by the reversal of charge point of the phosphatide, in a middle range of the pH — where the complex relations are much stronger — only flocculi are produced. The latter can however be made to sinter together by rolling between cover glass and microscope slide after which these rolls become eventually perfectly transparent and optically empty[2]. See fig. 29.

a   b   c
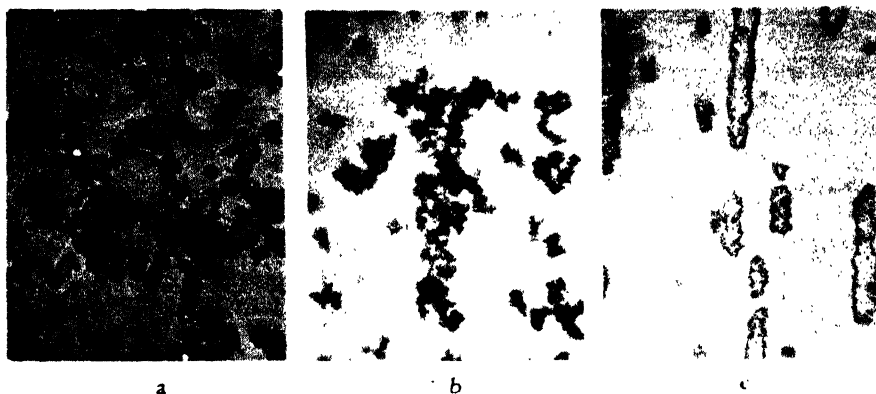
Fig. 29. Different morphological character of the complex coacervate gelatin (positive) : phosphatide (negative) at pH 4.5 and 3.5.
a. larger coacervate drops at pH 4.5.
b. flocculi at pH 3.5.
c. after pushing the cover glass backwards and forwards with the preparation depicted in b the flocculi sinter to rolls which gradually (for example, after one hour) become completely transparent and optically clear.

The flocculi therefore consisted of a very viscous coacervate. As was to be expected, on the addition of an indifferent salt these flocculi are first transformed into thoroughly liquid coacervate drops and at still higher salt concentrations the coacervation is suppressed. The added salt increases the water content by weakening the complex relations (p. 364), whereby the coacervate becomes less viscous. As a result rapid complete fusion occurs of the innumerable very small coacervate drops, which in the flocculi were only superficially fused together here and there.

q. Systems of higher order, preceding complex flocculation or complex coacervation

In some cases (e.g. positive proteins + negative phosphatides) complex flocculation is established rapidly at favourable mixing proportions of the colloid components but

---

[1] Association colloid!
[2] H. G. BUNGENBERG DE JONG and R. F. WESTERKAMP, *Bioch. Z.*, 234 (1931) 367.

at less favourable mixing proportions opalescent systems of comparative stability result, in which practically no flocculi can be seen on microscopic investigation [1].

On the addition of a little indifferent salt microscopically or macroscopically readily visible flocculation sets in rapidly.

The original opalescent systems are *"Apparent single colloid systems"*, which behave quite as hydrophobic sols (p. 234, Ch. VIII, § 1c). The "particles" of the sol are ultra-microscopic coacervate drops, which are capillary electrically charged on their surface by the complex component present in excess. Addition of salt suppresses the comparative stability of the "sol" and in this a salt is effective at concentrations which are the smaller the higher the valency of the oppositely charged ion. The ultra microscopic coacervate drops now begin to stick together into flocculi.

Only at higher salt concentrations the suppression then (after possibly also passing through a stage of readily visible coacervate drops) proceeds in the usual way according to the double valency rule.

The above described systems of higher order could be called *"coacervate sols"*.

Their relative stability is promoted by electrolyte-poverty of the medium and by unfavourable mixing proportions as well as by larger charge density of the colloid components. Through the charge density factor the coacervate sols obtained in the combination of a phosphatide with clupein are much more characteristic than in the combination with gelatin.

If the relative stability of the coacervate sol be removed, for example, by addition of electrolyte the system will flocculate. In the process clusters of very small coacervate drops are formed in the first instance which with sufficient fluidity of those drops may coalesce into larger drops or into a single coacervate layer.

Since however larger charge density (small equivalent weight), as is discussed below (§ 2r) results in general in a smaller water content of the complex coacervate, the ultramicroscopic coacervate drops in the combination with clupein are very viscous or possibly glass-like in nature, which greatly impedes the mutual fusion into larger drops.

Of course the primary formation of a "coacervate sol" precedes every rapidly occurring macroscopic coacervation. Through the above mentioned conditions not being fulfilled it is however unstable and is immediately destroyed by fusion into larger coacervate drops (compare p. 235-236).

r. Significance of the equivalent weight of the complex components

In the course of investigations on complex coacervation or complex flocculation specific differences were observed [2]: positive gelatin does indeed give complex coacervation with negative nucleate or with negative gum arabic, with negative agar only a slight opalescence is produced, while with negative soluble starch or negative glycogen complex coacervation and complex flocculation are entirely absent. The NaCl resistance (concentration of NaCl just sufficient for suppression) decreases strongly in the same series. It is greater in the combination with nucleate than with arabinate, while the combination with agar already loses its opalescence at very small NaCl concentrations. The combinations with starch or glycogen have a salt resistance equal to zero. Thus with the same constant positive complex component

[1] H. F. BUNGENBERG DE JONG and R. F. WESTERKAMP, *Bioch. Z.*, 234 (1931) 367.
[2] H. G. BUNGENBERG DE JONG and A. DE HAAN, *Bioch. Z.*, 263 (1933) 33.

(gelatin) the negative complex components can be arranged in a series with decreasing intensity of the complex relations.

Nucleate ⟩ arabinate ⟩ agar ⟩ amylum solubile, glycogen.

Now this is the order in which the equivalent weight increases from left to right (p. 270, Table 2).

It had already been observed earlier [1] in similar investigations regarding complex coacervation or flocculation of positive gelatin or of positive clupein with a number of negative sols of phosphatides of various origins that in both cases the order of the phosphatide sols with regard to the salt resistance is the same but this latter lies much higher in the combinations with clupein than in the combinations with gelatin.

The same holds for the combinations of these two proteins with the above mentioned series nucleate to amylum solubile. Thus we can in principle set up a similar series also for the positive complex components, in which series, if combined with the same negative component, the intensity of the complex relations decreases from left to right. This series (consisting here of only two members) is as follows:

clupein ⟩ gelatin.

Later measurements have established that the equivalent weight of positively charged clupein is in fact much smaller than that of positively charged gelatin [2]. The intensity of the complex relations each time at the most favourable mixing proportions and pH thus depends both on the negative and on the positive complex partner.

The complex relations are more intense the smaller the equivalent weight. One thing and another agree with the ideas already developed in § 2 o (p. 370). The complex is stronger the greater the number of bonds of salt-like nature per gram of colloid-colloid salt.

The magnitude of the equivalent weights of the two partners is indeed the most important factor for the complex coacervation. But not only differences in salt resistance but also other differences are governed by this.

Fig. 30 gives two coacervate volume curves at pH about 3.7; the one for an isohydric series of mixtures of 2% sols of gelatin (G) and Na nucleate (N), the other for 2% sols of gelatin (G) with Na arabinate (A) [3].
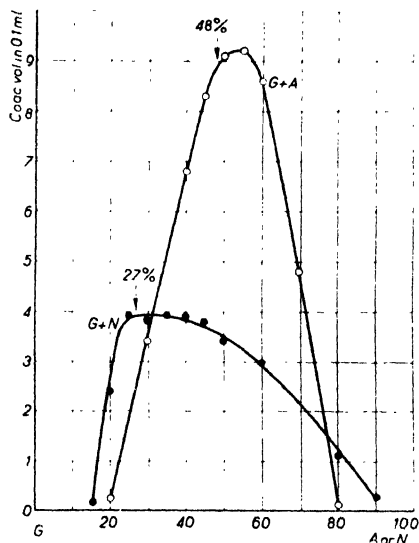
Fig. 30. Comparison of the complex coacervation of gelatin + nucleate with gelatin + arabinate (pH 3.7)
Ordinates: coacervate volume in 0.1 cc produced from 15 cc buffered colloid mixture. Abscissae: mixing proportion of the 2% sols expressed in % of arabinate (A) or nucleate (N). Arrows indicate the electrophoretically determined reversal of charge points.

[1] H. G. BUNGENBERG DE JONG and R. F. WESTERKAMP, *Bioch. Z.*, 234 (1931) 367.
[2] L. TEUNISSEN-VAN ZIJP, *Thesis*, Leiden, 1938.
[3] H. G. BUNGENBERG DE JONG and E. G. HOSKAM, *Proc. Koninkl. Nederland. Acad. Wetenschap.*, Amsterdam, 45 (1942) 585.

The reversal of charge points are indicated by arrows. We see that the equivalent mixing proportion lies much more towards the left for the combination G + N than that for the combination G + A. Since the equivalent weight of N is much smaller than that of A (p. 265 Table 1) this is also quite to be expected.

Further one sees that the maximum of the coacervate volume curve lies much lower for G + N than for G + A. This happens because the G + N coacervate is appreciably richer in colloid than the G + A coacervate (the G + N coacervate is also much more viscous). This relatively smaller water content of the G + N coacervate is also to be expected from the much stronger complex relations in this combination (§ 2o, p. 370).

The question whether still other factors beside the equivalent weight are important for the strength of the complex relations cannot at present be answered. It is possible that the nature of the ionised groups [1] plays a part but this factor is probably only of secondary importance compared with the equivalent weight.

s. Specific ion sequences in suppression by indifferent salts

As regards the influence of salts on complex coacervation or complex flocculation we have presented it as if only the valency of the cations or anions were of importance (the relevant valency rules have already been discussed in § 2 f, see p. 349 and 353).
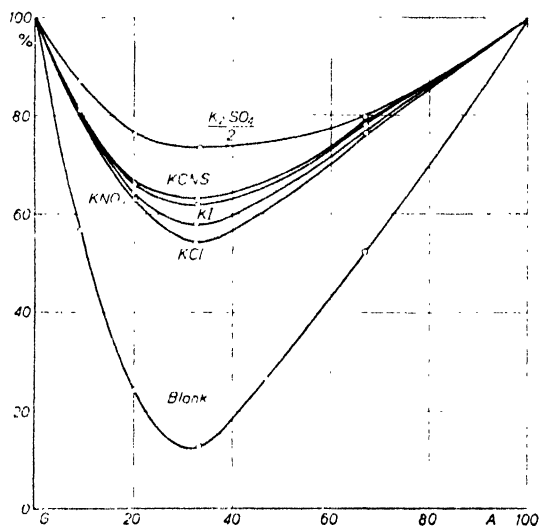
Fig. 31. Viscosimetric investigation of the suppressive action of salts of the type 1—1 (in which the anion is varied) on the complex coacervation of 0.67% gelatin and gum arabic sols at pH 4.0 (see caption of Fig. 9, p. 349). The lowest curve relates to mixtures without added salt. The remaining ones to mixtures with 13.3 m. eq. per l. added salt.
The return to additive behaviour increases in the order: Cl — NO$_3$ — I — CNS (specific differences between the various monovalent anions).
K$_2$SO$_4$ belonging to the salt type 1—2 (valency rule of the anions) has a stronger suppressive action than these salts of the type 1—1.

In the study of the complex coacervate gelatin — gum arabic it was seen quite early on that this is only a first approximation. Here indeed LiCl, NaCl and KCl (monovalent cations) have almost the same action as one another, similarly MgCl$_2$, SrCl$_2$ and BaCl$_2$ (divalent cations) but on comparing KCl, KI, KNO$_3$ and KCNS

---

[1] Compare the order phosphate > sulphate > carboxyl in the flocculability of acidoids with large organic cations (see note 2 on p. 405). Since proteins also appear here as organic cations the above series might be the sequence of the strength of the individual salt bond in the combinations of positive proteins with phosphate, sulphate and carboxyl colloids.

relative differences already clearly appeared between the above mentioned monovalent anions. See Fig. 31.

In general it may be foreseen that specific differences can occur between ions of the same valency and then these ions will have to arrange themselves in the sequences in which they appear in the reversal of charge spectra of the colloid components.

In the ion spectrum of the positive proteins the reversal of charge concentration increases in the order I $<$ Br $<$ Cl, similarly in the order CNS $<$ NO$_3$ (p. 299, Ch. IX, § 3b), in that of the carboxyl and sulphate colloides the reversal of charge concentration increases in the order K $<$ Na $<$ Li (p. 284 and 285, Ch IX, § 2d and 2e), on the other hand in that of the phosphate colloids: Li $<$ Na $<$ K (p. 280, Ch. IX, § 2c). From this one must expect that the necessary salt concentrations for the suppression of the complex relations must *increase as follows from left to right* for the salts of type 1—1:

positive protein + sulphate colloid
positive protein + carboxyl colloid

$$\left\{ \begin{array}{l} \text{KCNS} < \text{KNO}_3 \\ \text{KJ} \quad \text{KBr} \quad \text{KCl} \\ \text{KCl} \quad \text{NaCl} \quad \text{LiCl} \end{array} \right.$$

positive protein + phosphate colloid

$$\left\{ \begin{array}{l} \text{KCNS} \quad \text{KNO}_3 \\ \text{KJ} \quad \text{KBr} \quad \text{KCl} \\ \text{LiCl} < \text{NaCl} < \text{KCl} \end{array} \right.$$

An orientating investigation in this direction with several complex combinations for the suppression with these and other salts of the type 1 — 1 or 2 — 1 gives in general a confirmation of these expectations[1]. In 11 of the 12 different complex combinations investigated on this point the expected order of the alkali cations occurred. As regards the orders to be expected for the anions these occurred in 7 of the 10 complex combinations investigated. In the three discrepant combinations: clupein (positive) + phosphatide (negative), gelatin (positive) + phosphatide (negative) and gelatin (positive) + nucleate (negative) exactly the opposite orders occurred:

$$\text{KCl} < \text{KBr} < \text{KI and KNO}_3 < \text{KCNS.}$$
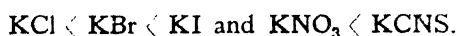
We think we must give the following explanation[2] for this reversal: There is no reason for assuming that these anions should influence the positive complex component in these complexes otherwise than that which corresponds to their order in the ion spectrum, therefore the reversal must be sought in this case in a reciprocal action with the negative complex component. This negative component is here not exclusively a carrier of negatieve ionised groups. The negative charge of the phosphatide comes about through an acid admixture (probably phosphatidic acid) in the much larger amount of phosphatide amphoions which together form the kinetic units (association colloid). In the dicomplex coacervation or flocculation the "phosphatide" reacts as a colloid anion precisely through this excess of negative charge over the positive charges also present. It therefore reacts with the algebraic sum of its negative and positive charges. Nevertheless it must be borne in mind that circumstances can arise in which the positive charges themselves also begin to play a part[3]. We

[1] O. BANK and E. G. HOSKAM, *Protoplasma* 34 (1940) 188; see also sequences found earlier H. G. BUNGENBERG DE JONG and R. F. WESTERKAMP, *Bioch. Z.* 234 (1931) 367.
[2] The authors give a different explanation, which seems to us to be too complicated.
[3] This is for example, the case in the tricomplex systems. See § 6, p. 415.

now believe these to be present here. For these positive charges naturally there holds just as much the same affinity order as for albumin cations: thus I ⟩ Br ⟩ Cl and CNS ⟩ NO₃. A suppressive salt can in the complex combination positive protein + negative phosphatide therefore also react with its anions on the negative component, the result of which is however that by binding to the positive groups of the phosphatide the effective negative charge of the phosphatide increases[1]. This would in itself result in an increase of the salt resistance.

Since this binding of anions increases in the order Cl ⟨ Br ⟨ I and NO₃ ⟨ CNS the order of the suppression must as a consequence of this be the one experimentally found. This therefore occurs because this action of the anions on the phosphatide in the final result predominates over the action of the anions on the positive protein component.

Similar considerations hold for the negative component nucleate. Acid functions as well as basic functions are present in the molecule but the acid ones far and away predominate. In the acid medium in which the complex combinations are investigated one must however take account of positive charges. If the attachment of anions to these is strong then this can lead to reversal of the anion series.

This is obviously the case in the complex combination gelatin + nucleate. On the other hand the anion order in the combination clupein + nucleate is normal, that is to say, the influence of the anions on the protein component predominates here. Now with this combination the medium is not so acid as in the case of gelatin, so that the anion influence on the nucleate remaining in the background is explicable.

t. Complex coacervation in the presence of three colloid components

The mixing proportions of three isohydric sols can be represented in the so-called ternary triangle, the corners of which represent these sols themselves. In choosing equal colloid concentrations, such a diagram allows to represent the mixing proportions of the three colloids. This triangle is of course only a simplified phase theoretical isotherm, as the fourth component present — the water — is left out of consideration.

In the investigation of systems with three complex components two different cases can occur. The first case occurs in the combination gelatin *(G)* + ichthyocoll *(I)* + gum arabic *(A)*. If we choose the pH such that the two proteins are positively charged, the gum arabic has not yet lost its negative charge, then complex coacervation occurs in the two "binary" combinations *G + A* and *I + A* in a certain range of mixing proportions. Now it has been found that in mixtures which contain *G*, *I* and *A* simultaneously always only one complex coacervate can occur. It follows that the two "binary" complex coacervates *G + A* and *I + A* are miscible in all proportions. This case therefore opens up no new aspects and was not further investigated quantitatively. Fig. 32 indicates schematically what may be expected as a result. The "binary" mixtures *G + A* lie on the side *GA* of the triangle. Let *a* and *b* represent the limits of the coacervation range in this combination and point *r* the reversal of charge point in this coacervation range.

---

[1] In agreement with this one finds that at the reversal of charge of lecithin with Ca salts the reversal of charge concentration rises from left to right in the order:
    CaCl₂, CaBr₂, CaI₂ and Ca(NO₃)₂ ⟨ Ca(CNS)₂ compare P. H. TEUNISSEN and H. G. BUNGENBERG DE JONG, *Kolloid Beih.*, 48 (1938) 33, see p. 50, Fig. 12.

The points c and d and point r on the side $A\,I$ of the triangle have the same significance for the "binary" isohydric series of mixtures of I with A We may then expect that the coacervation phenomena in mixtures of the three colloids can be appreciated when the course of three curves is known, namely the curve which connects a with c, the curve which connects the two reversal of charge points r and the curve which connects b and d.

Fig. 32.   Scheme showing in principle the complex coacervation of equally concentrated isohydric gelatin (G), ichthyocoll (I) and arabinate (A) sols. $a\,b\,d\,c\,a$ the region in which complexcoacervation takes place. The dotted line, which joins the two reversal of charge points r on the sides GA and IA, gives the reversal of charge points inside the region $a\,b\,d\,c\,a$.
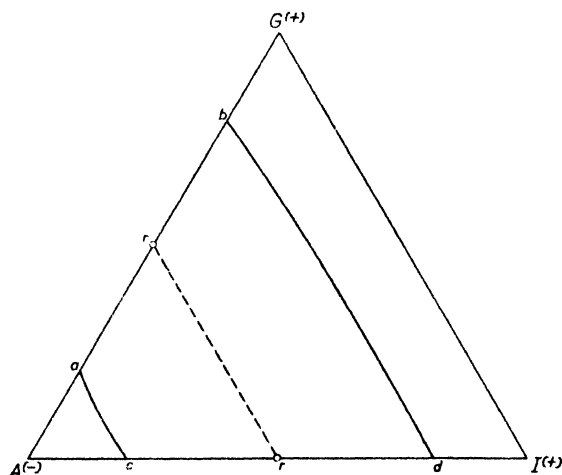
The plane of the triangle is divided by the dotted reversal of charge curve into two halves. The part between corner A and the dotted curve includes all the systems which are negatively charged, the part between the dotted curve and the side $G\,I$ all the systems which are charged positive. The curves ac and bd divide the plane of the triangle into three parts, the middle one of which $(a\,b\,d\,c\,a)$ includes the mixtures in which complex coacervation takes place. The reversal of charge curve is situated in this field. In the two other fields $(AacA$ and $bGIdb)$ complex coacervation does not occur.

The second example of complex coacervation in mixtures of three colloid components is much more interesting and has been investigated in somewhat greater detail[1]. The colloids in question are: gelatin (G), arabinate (A) and nucleate (N). The pH ($\pm$ 3.7) was so chosen that G was positive, A and N were negative. Fig. 33 reproduces the results for approximately isohydric 1.3% sols at 40°.

This figure for the greater part corresponds with Fig. 32, so that an explanation need not be repeated. The curve, which joins the reversal of charge points on the sides of the triangle (u $=$ o) was found to be a straight line (p. 328, Ch. IX, § 6d) and the curves ab and cd were slightly curved.

The point of difference with Fig. 32 is however that in the central part of the coacervate field $abdca$ a new phenomenon occurs. Here lies a field of mixing proportions enclosed by a curve, in which two coexisting coacervates occur. Both coacervates are complex coacervates, which contain all three colloids, but the one is rich in A and poor in N, the other is on the contrary rich in N and poor in A.

[1] H. G. BUNGENBERG DE JONG and A. DE HAAN, *Bioch. Z.*, 263 (1933) 33; H. G. BUNGENBERG DE JONG and E. G. HOSKAM, *Proc. Koninkl. Nederland. Akad. Wetenschap.*, Amsterdam, 45 (1942) 387; H. G. BUNGENBERG DE JONG, *Proc. Koninkl. Nederland. Akad. Wetenschap.*, Amsterdam, 45 (1942) 393, see p. 394—396.

We can therefore bring out the differences between the systems $G + I + A$ and $G + A + N$ by saying that in the first the two "binary" complex coacervates $(G + A$ and $I + A)$ are miscible in all proportions and that in the second system the two complex coacervates $(G + A$ and $G + N)$ are only partially miscible. Now
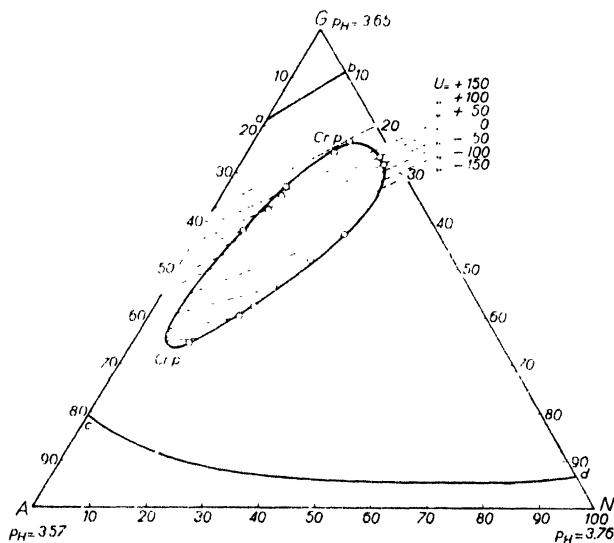
Fig. 33. Complex coacervation in mixtures of gelatin (G), Na arabinate (A) and Na nucleate (N) (pH ± 3.7) Inside the closed ellipse-shaped region two coexisting complex coacervates make their appearence. Outside this region but inside the region $a\ b\ d\ c\ a$ only one complex coacervate appears. Dotted lines going from the side AG to the side GN give mixtures of constant electrophoretic velocity U (in arbitrary units). One of them, namely that for $U = 0$, is thus comparable with the dotted line in Fig. 32. The lines of constant U must, between their intersections with the ellipse, also represent tie-lines[1] between coexisting $G + A + n$ and $G + N + a$ coacervates.

Some dotted lines lower in the ellipse give the probable position of the tie-lines. The $G + A + n$ coacervates lie on the left hand arc of the ellipse, the $G + N + a$ coacervates on the right hand arc of the ellipse between the two „critical points".

for an explanation it is also necessary to bear in mind that the reversal of charge line also runs through this central field, in which two coacervates as well as equilibrium liquid are produced. Now we already know that the complex relations are a maximum close to the reversal of charge point (p. 345, § 2 e and p. 357, § 2h). At this point therefore the specific differences between the two binary complex coacervates will be able to manifest themselves most strongly.

Now in the pH range 2 - 5 G and I resemble each other relatively greatly in their properties[2] and thus the maximum complex relations will differ but little in the complex coacervates $G + A$ and $I + A$. This is obviously favourable for their complete miscibility.

On the contrary there is an appreciable difference in equivalent weight between N and A. The complex coacervates $G + N$ and $G + A$ do differ much from one another and there is every indication that the maximum complex relations (at the same pH) are considerably stronger in the first than in the second (p. 374, § 2 r). It is

[1] Properly spoken these line are projections of tie lines occurring in the tetrahedral isotherm Water$+G+A+N$. The fourth component, water, has however been left out of consideration in the triangles of figs. 32 and 33, which represent only the colloid compositions of the mixtures.

[2] e.g. shape of the viscosity-pH curve. See p. 208, Fig. 20.

obviously this circumstance which leads to the partial miscibility of these complex coacervates.

As a result of this partial miscibility the two coexisting coacervates contain all three complex components. The one besides gelatin contains mainly A together with a little N; the other besides gelatin contains mainly N together with a little A. To distinguish them these coacervates can be denoted by the symbols $G + A + n$ and $G + N + a$ respectively, whereby the components mainly present are therefore denoted by capital letters. In Chapter XI, § 1f-h (p. 438) we discuss further the mutual wetting properties of these coexisting coacervates, as a result of which drops of $G + A + n$ and of $G + N + a$ unite to form composite coacervate drops.

For further details (among others the influence of the pH, position of the lines which connect coexisting coacervates) reference may be made to the literature quoted.

If one brings about complex coacervation in a mixture of gelatin, gum arabic and amylum solubile, then one can readily show that all the amylum solubile is to be found in the equilibrium liquid and the complex coacervate contains no amylum solubile; thus the coacervate is exclusively $G + A$. It is possible we have here an extreme case of that described above. Negative amylum solubile in the properties which are in the first place determinative for complex coacervation stands very far from gum arabic. Its equivalent weight is very large (see p. 270 Table 2) and already it no longer forms complex coacervates with G (salt resistance zero). It is understandable that it also cannot compete with A for the G in the complex coacervate, so that all the amylum solubile must accumulate in the equilibrium liquid.

u. Complex Gels

If one places a thin plate of gel, obtained by gelation of a gelatin sol in an appropriate mould, in dilute acetic acid, then this gel plate swells very considerably.

If however one has prepared such a gel plate by cooling a suitable mixture of gelatin sol and gum arabic sol, then this swelling does not take place in dilute acetic acid and furthermore a considerable turbidity occurs in the gel.

After the acetic acid has diffused in, the gelatin cohering to a gel structure becomes positive and now complex relations appear with the negative gum arabic embedded in between the gel structure, with the result of suppression of the tendency to swell and of turbidity[1].

We can call such a gel system a *complex gel* (p. 335, § 1) and in the case in question, one belonging to the variant colloid cation + colloid anion.

We must thus expect that as in the analogous complex coacervation gelatin + gum arabic, the intensity of the complex relations (as a measure of which we can choose the degree of turbidity) depends on the mixing proportion of the two colloids in the gel and at constant mixing proportion on the pH. Both are true. Further salts must weaken the complex relations, that is to say cause the turbidity to decrease and the ability to swell to reappear. This is in fact also found and as was to be expected the double valency rule makes its appearance. For the same concentration (in m. eq. per 1) the turbidity is suppressed more strongly the greater the valency of the cation, the valency of the anion remaining constant; similarly the greater the valency of the anion, the valency of the cation remaining constant.

The suppressive action therefore decreases in the following series of salts from left to right:

$$3 - 1 \rangle 2 - 1 \rangle 1 - 1 \text{ and } 1 - 3 \rangle 1 - 2 \rangle 1 - 1.$$
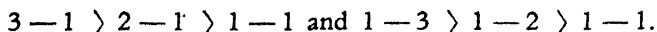
[1] H. G. BUNGENBERG DE JONG and H. J. C. SENGERS, *Rec. trav. chim.*, 53 (1934) 171.

Complications occur in the investigation of the suppressive action of $K_3Fe(CN)_6$ and $K_4Fe(CN)_6$ which indicate that the complex relations between the positive gelatin and the negative gum arabic are indeed suppressed but that new complex relations between positive gelatin and the 3 and 4-valent anions in question are produced so that with these salts the original complex gel variant colloid cation + colloid anion is transformed into a complex gel variant colloid cation + micro anion.

In the complex gels discussed only one component, the gelatin, is present as a gel structure, the other component (arabinate anion or Fe $(CN)_6''''$) is in principle mobile.

There exist however also complex gels in which the two components form an intertwined gel structure which structures have struck up complex relations between one another (for example gelated mixture of gelatin and agar sols, placed in an acid medium).

Complex gels can also be obtained in another — really much more natural — way namely by cooling complex coacervates containing gelatin. If for example the separation into layers has taken place at 40° to a clear equilibrium liquid and clear coacervate layer and one subsequently cools the tube to room temperature both layers become turbid. The coacervate layer becomes solid thereby and on microscopic examination the turbidity is seen to be caused by a large number of small vacuoles in the gel. The equilibrium liquid becomes turbid because some fresh coacervate separates out of it in small drops. These drops also gelate and stick to each other forming loosely built flakes.

In the gelation of the complex coacervate a change of the "solubility" of the coacervate has taken place just preceding the gelation, indeed the coacervate becomes poorer in water (vacuoles in the gel) and colloids (fresh coacervate) separate out from the equilibrium liquid [1].

For the investigation of these complex gels one can however start with advantage from gelated complex coacervate drops of microscopic dimensions, the preparation of which we leave out of consideration here. These gelated coacervate drops readily stick to a glass surface and one can now make any desired liquid, for example a very dilute buffer of the desired pH, flow past these drops in a suitably constructed cuvette and now measure the changes in diameter. Their very small dimensions are seen to bring with them the extraordinarily great advantage that the swelling equilibria are established very rapidly (for example within 5 minutes) and it thereby becomes possible to investigate whether the diameter changes are reversible by alternately passing two liquids through the cuvette. Such an investigation on gelated complex coacervate drops of gelatin + gum arabic [2] has contributed essentially to deepen our knowledge of complex gels.

On varying the pH the diameter of the gelated drops goes through a minimum which lies very near that pH (3.7) at which the complex coacervation is an optimum at the given mixing proportion of gelatin and gum arabic.

If now one works further at constant pH (3.7) and one investigates the influence of an added indifferent salt, all salts are seen to cause the diameter to increase. These changes are reversible provided the salt concentration is not too great. Furthermore

---

[1] The same phenomena (vacuolation, formation of flakes) occur also in cooling a suspension of coacervate drops in their equilibrium liquid. See p. 449, Fig. 16.

[2] H. G. BUNGENBERG DE JONG and J. M. F. LANDSMEER, *Rec. trav. chim.*, 65, (1946) 606.

the double valency rule (see p. 350) makes its appearence in this connection (see Fig. 34). The increase in diameter by the salts increases in the following series from right to left:

$3 — 1 \rangle 2 — 1 \rangle 1 — 1$

and

$1 — 3 \rangle 1 — 2 \rangle 1 — 1$

Summarising we thus see that the same factors which are determinative for complex coacervates in influencing the intensity of the complex relations (the pH at constant mixing proportion of the colloids; suppressive action of salts) also hold with these complex gels and manifest themselves in reversible changes of the amount of swelling water (compare p. 450, Fig. 18, in Ch. XI). Similar changes in the water content occur with complex coacervates (p. 368, § 2n).
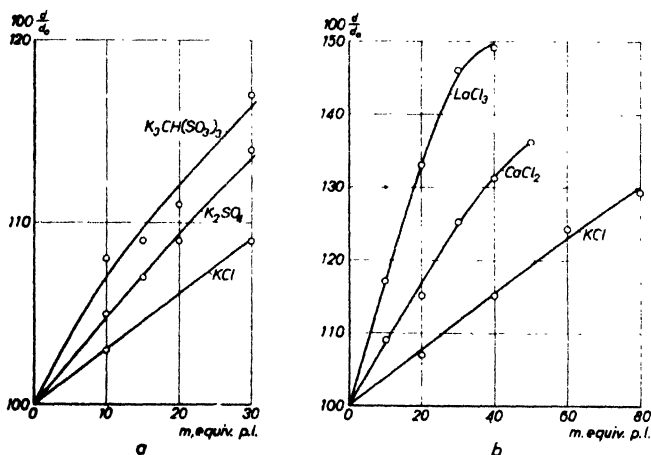
Fig. 34. Swelling of the gelatinised complex coacervate drops A+G by salts.
Ordinates: diameter of the complex gel drops in % of the initial value before salt addition.
Abscissae: Salt concentration in m eq. per l.

If however one chooses the concentration of a salt high enough then irreversible changes of diameter occur either immediately or after a few alternations of the liquid[1].

In this a complete reversal of the usual behaviour occurs. While normally addition of salt leads to swelling with a complex gel, a shrinkage then occurs and after a few further alternations of liquid this inverse behaviour has also become reversible.

If one subsequently investigates the diameter changes as f(pH) it is then found that the swelling minimum no longer lies at the original value of the complex gel (e.g. 3.7) but it is shifted appreciably towards higher pH and, if the salt treatment had been sufficiently effective, now coincides with the isoelectric point of the gelatin used.

That is to say the original complex gel has lost all its gum arabic and has been transformed into a gelatin gel. This makes the above mentioned inversion understandable at the same time. In fact salts no longer exert a swelling action but rather a shrinking one on a gel consisting exclusively of positive gelatin at pH 3.7[2]. We thus arrive at the following picture: In the original complex gel the complex partner gum arabic which is in principle mobile is bound to the gelatin gel structure by intense complex relations. It is not then delivered to the liquid flowing past. If now the complex relations are diminished in intensity by indifferent salts, then at first the

[1] H. G. Bungenberg de Jong and J. M. F. Landsmeer, *Proc. Koninkl. Nederland. Akad. Wetenschap.*, Amsterdam, 51 (1948) 137 and 295.
[2] H. G. Bungenberg de Jong and J. Ph. Hennemann, *Kolloid Beih.*, 36 (1932) 123.

same conditions still practically hold, i.e. the swelling by indifferent salt is reversible.

At sufficiently high salt concentrations however the complex relations are completely suppressed and the mobile component, the gum arabic, can now pass out rather rapidly by diffusion and is removed for good along with the liquid flowing past. The complex gel is thus transformed into a gel containing exclusively gelatin.

Conversely one can now also obtain a complex from this again by passing a very dilute gum arabic sol of pH $3\cdot7$ over the gelatin spheres. They shrink then once more in their dimensions and again assume the properties of the complex gels and swell reversibly with salts. The mobile complex partner thus can not only leave the gel on suppression of the complex relations but can just as well be taken up in the positive gelatin gel when it is offered from without and the conditions for the contracting of complex relations is favourable (pH $3\cdot7$ in the absence of salt).

Wheat gluten — the plastic and elastic mass which is obtained by washing out wheat flour with water — has for a long time been considered as a mixture of two proteins containing water or better of two groups of protein components which to distinguish them are denoted by the terms *gliadin* and *glutenin*.

H. L. BUNGENBERG DE JONG[1] showed with purified preparations of gliadin and glutenin that complexes are formed between them in the pH range between the two isoelectric points (where gliadin is charged positive and glutenin negative). The intensity of the interaction depends in the usual way on the mixing proportion and the pH. From this it is very probable that wheat gluten is not a mixture but a complex system of gliadin and glutenin at any rate in the pH region between the two isoelectric points.

§ 3.   DICOMPLEX SYSTEMS II. THE VARIANTS COLLOID CATION . MICRO ANION AND MICRO CATION ·i COLLOID ANION

a.   The coacervation of gum arabic with hexol nitrate as an example of the variant micro cation + colloid anion

We have already discussed in previous chapters coacervation or flocculation by added indifferent salts (for example sulphates) in which this separation of the colloid usually occurs only at relatively high concentrations (p. 202, 226, 252). The colloid is here expelled from the solution because the ions of the added salt bind the water in the order of the lyotropic series.

In this "salting out" there are no indications that the ionised groups of the colloid play an active part in the production of the coacervate or flocculation.

In the flocculations or coacervations which are discussed in this paragraph indications of this kind are indeed very clearly present. The required concentrations of the neutral salt are here not necessarily very high, on the contrary they are frequently relatively low. In so far as lyotropic influences are also found in the coacervations or flocculations now in question, they are not due to a binding of the water as mentioned above, but to the fact that the interaction between the ionised groups of the colloid and the ions of the salt is influenced by the same factors (radius, polarisability) which determined the order of the ions in the lyotropic series.

[1] H. L. BUNGENBERG DE JONG, *J. Soc. Chem. Ind.*, 52 (1933) 391.

As an example in which the above mentioned concentrations are even very small, we shall discuss the coacervation of gum arabic with hexol nitrate[1] (a complex salt with a hexavalent cation[2]) as an introduction to what follows later. If one adds hexol nitrate solution to a gum arabic sol, a turbidity is produced on exceeding a certain small quantity and is found on microscopic examination to consist of intensely brown coloured liquid drops (see Fig. 35) (Hexol nitrate has a brown colour in solution).
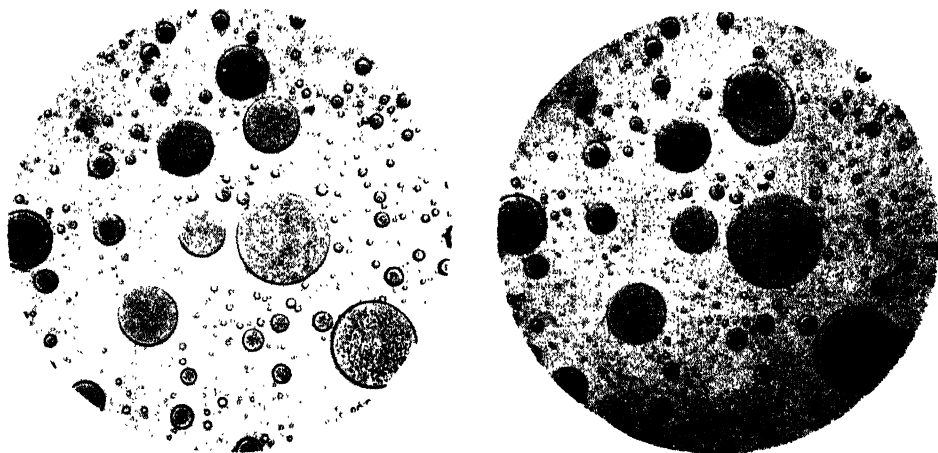
Fig. 35.   Coacervate of gum arabic with hexol nitrate.
The coacervate drops were weakly vacuolised.
Their thoroughly liquid nature appears from comparison of the two microphotos. On the right one sees one large drop lying at the place where on the left there are two small ones. Between making the two exposures these drops have encountered one another and coalesced.

The concentration required for coacervation just to set in is now practically proportional to the gum arabic concentration (see the straight line A in Fig. 36b on p. 386), showing that the word concentration has not the significance here of free hexol ions present in the sol but these latter are bound to the gum arabic.

The coacervate drops are now electrophoretically negatively charged but on adding more hexol nitrate they reach the reversal of charge point (see straight line B in Fig. 36b) and at still greater concentrations they become electrophoretically positive. The same statement as above also holds for the hexol nitrate "concentrations" for reaching the reversal of charge point: the added hexol ions — except in extremely dilute gum arabic sols — are practically all bound to the gum arabic.

We have already discussed this in more detail in Chapter IX, § 1b, p. 262, from which it was seen that the true reversal of charge concentration is exceedingly small in this case (of the order of $10^{-6}N$) and is determined by the intersection of the straight line B with the ordinate axis.

From the above it is already seen that for the coacervation of gum arabic with hexol nitrate what matters is the ratio of the total amount of gum arabic present

[1] H. G. BUNGENBERG DE JONG and J. LENS, Biochem Z., 235 (1931) 185.
[2] $[Co\{(OH)_2 Co(NH_2CH_2—CH_2NH_2)_2\}_3[(NO_3)_6$

and the total amount of hexol ions present, independently of the amount of water present.

When we add a fairly large excess of hexol nitrate (order of 0.1 N) the positive coacervate finally also goes into solution once more (compare C in Fig. 36b although the intersection with the ordinate axis is here drawn very much too low in proportion).
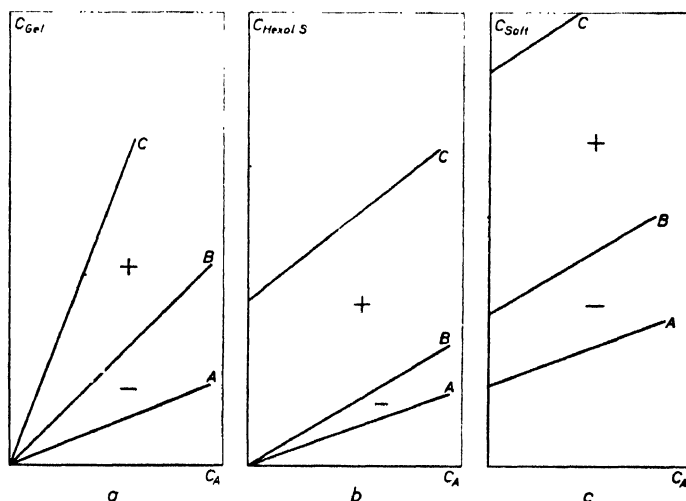
Fig. 36. Reversal of charge, beginning and end of the coacervation or flocculation range as a function of the colloid concentration C_A.

a: in the complex coacervation in the narrow sense. (see p. 338, § 2).
b: in the coacervation of gum arabic with hexol nitrate (see also p. 388).
c: in the coacervation or flocculation of gum arabic or an other acidoid with lower valent cations (see p. 392).
A and C beginning and end of the coacervation, B electrophoretic reversal of charge. In the coacervation region situated between A and C the coacervate drops are negative between A and B, positive between B and C.
In a, for the sake of comparison, the complex coacervation gelatin (positive) + gum arabic (negative) is considered as coacervation of the gum arabic by the salt gelatin chloride. When one considers in this case small concentrations of gum arabic the curves A, B and C are straight lines which pass almost through the origin.

Here we see already a certain analogy to the ordinary complex coacervation, arabinate (negative) + gelatin (positive): in this latter also coacervation only occurs in a certain range of mixing proportions and reversal of charge occurs in this coacervation range. According to Fig. 46 (p. 327) the reversal of charge point also lies at a particular mixing proportion of the two colloids, independently of the amount of water present.

The similarity also appears to continue in all kinds of other properties. Thus the coacervate drops obtained from gum arabic and hexol nitrate placed in an electric field also exhibit the *deformation and disintegration phenomena* previously discussed (see p. 347-348) and as concerns the latter these are with positively charged drops again the mirror image of those with negatively charged drops.

Furthermore added indifferent salts suppress the coacervation whereby the so-called *double valency rule* (see p. 350) again appears:

$$3 - 1 \ \rangle \ 2 - 1 \ \rangle \ 1 - 1 \text{ and } 1 - 4 \ \rangle \ 1 - 3 \ \rangle \ 1 - 2 \rangle \ 1 - 1$$

See Fig. 37 in which as the working point for the study of the salt suppressions the top of the turbidity curve with hexol nitrate (very near the reversal of charge point)

Fig. 37. Coacervation of gum arabic sol with hexol nitrate (left) and suppression of it by added salts (right).
The working point for the right-hand graph is indicated in the left-hand one by an arrow.
Ordinates: turbidity.
Abscissae: logarithms of the gross hexol nitrate concentration (left) and of the salt concentrations (right), both in eq. per l.
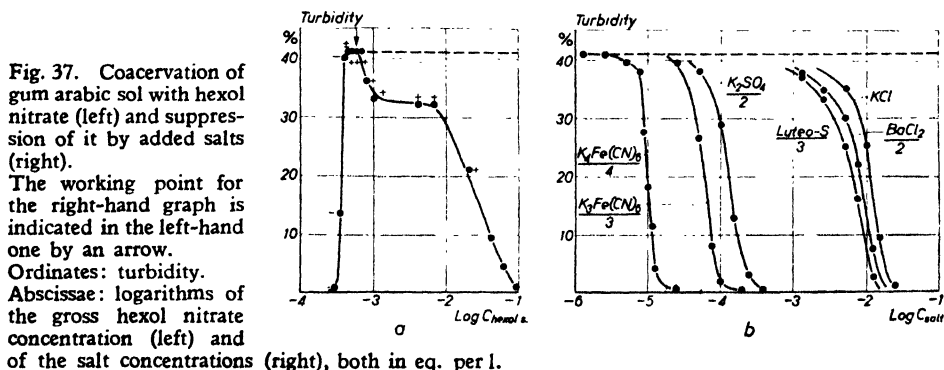
was taken. In addition the so-called *continuous valency rule* (see p. 352) was again encountered in this case for the influence of added neutral salts on the electrophoretic velocity of the coacervate drops[1]:

(relative $3 - 1 \dots 2 - 1 \dots 1 - 1 \dots 1 - 2 \dots 1 - 3 \dots 1 - 4$ (relative positivation)   negativation)

See Fig. 38 in which the upper graph relates to a negative coacervate, the lower graph to a positive coacervate obtained with rather more hexol nitrate.

Summarising we can say that the coacervation of gum arabic with hexol nitrate has all essential features in common with the complex coacervation, colloid cation + colloid anion.

Obviously an appropriately chosen micro cation — here the 6-valent hexol ion — can take the place of the colloid cation without changing the nature of the typical complex relations characteristic of the first mentioned variant.

All the arguments in the previous section (§ 2) will therefore also be applicable here. For example, one should bear in mind that in the Phase Theory the system Ca arabinate + hexol nitrate + H_2O does not form a ternary but at least a quaternary system (see p. 367) as the two salts have no ion in common.

It is thus again to be expected that quantitative results will still depend on the colloid concentration, and that, for example, at higher colloid concentrations the coacervate with hexol nitrate will be richer in water than at lower concentrations because the $Ca(NO_3)_2$ (formed from the non-essential ions) is to be found in greater concentration in the whole system and this will in principle act suppressive like any added salt.

To obtain as simple results as possible (for example the linear relations in Fig 36)

---

[1] H. G. BUNGENBERG DE JONG and J. L. L. F. HARTKAMP, *Rec. trav. chim.*, 53 (1934) 622.

it is thus essential to keep this $Ca(NO_3)_2$ concentration small, which is attained by always working with low colloid concentrations (see e.g. p. 263, Fig. 2).

In the above statements the emphasis was especially laid on the analogy with the complex coacervation, gum arabic (negative) + gelatin (positive). Meanwhile there is a striking point of difference. When one draws a graph valid for the complex coacervation (at least at sufficiently small concentrations of the isohydric sols) similar to Fig. 36 b, the curve C in it must begin practically at the corner (see Fig. 36 a). The sectors for the negatively charged coacervates (between A and B) and for the positive coacervates (between C and D) are here of the same order of magnitude.

Replacement of the gelatin cation by the hexol cation shifts the curve C very far upwards, as a result of which the region of the positive coacervates in Fig. 36 b (bounded by C, the ordinate axis and B) is extraordinarily much greater than the region of the negative coacervates (between A and B). Compare the asymmetric position of the reversal of charge point (far towards the left) in the coacervation range in Fig. 37a. The mechanism by which the upper coacervation limit (C) is reached with excess of hexol nitrate, is thus in this case quantitatively different from that with gelatin chloride.

In the latter case one can attribute the suppression to a progressive attachment of gelatin cations on arabinate anions, whereby the solubility of the coacervate (which is a minimum at the equivalent attached quantity) finally increases considerably.

In view of the large concentration of hexol nitrate needed, the suppression here rather bears the character of a suppression of the coacervation as a consequence of increase in concentration of the free ions present in the medium.

Fig. 38. Influence of added salts on the electrophoretic velocity of negatively (upper) or positively (lower) charged coacervate drops of Na-arabinate with hexol nitrate.
Ordinates: electrophoretic velocity in arbitrary units.
Abscissae: logarithms of the salt concentrations (in eq. per l.).
Horizontal dotted lines: electrophoretic velocity without added salt.

Since however the hexol ion itself brings about coacervation it must therefore in particular be the accompanying NO_3 ion which causes the suppression of the coacervation at sufficiently high concentrations.

The various facts would thus indicate that while progressive attachment of gelatin cations sufficient for suppression is indeed possible, this does not happen to a sufficient extent with increase of the hexol nitrate concentration on exceeding the reversal of charge point, as a result of which the suppression finally occurring assumes rather the character of a suppression by a salt (shielding).

This impeded attachment can be conceived as the consequence of the replacement of a cation (gelatin cation) which attaches itself with very great affinity by a cation (hexol cation) which attaches itself with less great affinity. On replacing the hexol cation by cations of lower valency, which attach themselves with still smaller affinity, the changes, which in the transition from Fig. 36a →fig. 36b still only show themselves in curve C, now are also visible in a similar displacement of the curves B and A, see Fig. 36c, that is to say, the practical proportionality between gross reversal of charge concentration or gross concentration at which coacervation or flocculation occurs and colloid concentration is lost.

For reversal of charge, coacervation or flocculation concentrations are now necessary, which, if curves A and B are displaced sufficiently far, gradually become practically independent of the colloid concentrations (at least if these latter are small).

b.   The coacervation of positive ichthyocoll sol with $K_3Fe(CN)_6$ as an example of the variant colloid cation + micro anion

Ichthyocoll is an isostable protein, which — apart from a different position of the isoelectric point — is in the main similar to gelatin in properties. Here as in the case of gelatin a simple behaviour is only present at temperatures (for example 37°) at which no gelation processes, whether slow or rapid, occur.

Strongly positively charged sols — for example at pH about 3 — give coacervation at 37° with $K_3Fe(CN)_6$ (Fig. 39). The coacervate drops, produced with little $K_3Fe(CN)_6$, are seen to be electrophoretically positive, those with more, on the other hand electrophoretically negative. The coacervate drops again show the typical disintegration phenomena in an electric field (see p. 347-348).

Fig. 40 reproduces turbidity measurements at pH 3.92 on 0.04% ichthyocoll sols. The left-hand figure relates to coacervation with $K_3Fe(CN)_6$ as a function of the concentration of the salt. The coacervation already begins at concentrations lower than 1 m. eq. per l (log C = 0.00—3), passes through a maximum at about 5 m. eq. per l (log. C = 0.70 —3) to decrease again later considerably at greater concentrations.

Fig. 39.   Coacervation of positive Ichthyocoll sol with $K_3Fe(CN)_6$. 48 × linear. The coacervate drops were weakly vacuolised.

The reversal of charge point also lies near the maximum of the turbidity curve.

The two graphs on the right of Fig. 40 relate to the influence cf extra added indifferent salts on the coacervation with a constant $K_3Fe(CN)_6$ concentration of 2 m. eq. per 1. The horizontal dotted line in the three graphs gives the turbidity which is produced by this in the absence of extra added indifferent salt. For the present paying attention only to the full curves in the right graphs we observe that all the salts used suppress the coacervation and indeed the double valency rule appears in it:

hexol nitrate $\rangle$ luteocobalt chloride $\rangle$ BaCl$_2$ $\rangle$ KCl and K$_2$SO$_4$ $\rangle$ KCl

or in salt symbols:

$$6 — 1 \ \rangle \ 3 — 1 \ \rangle \ 2 — 1 \ \rangle \ 1 — 1 \quad \text{and} \quad 1 — 2 \ \rangle \ 1 — 1.$$

Compared with the dicomplex coacervation, arabinate ion + hexol cation, there is nevertheless a point of difference. In the latter the "concentration" of the hexol nitrate required just to induce coacervation or to reach the reversal of charge
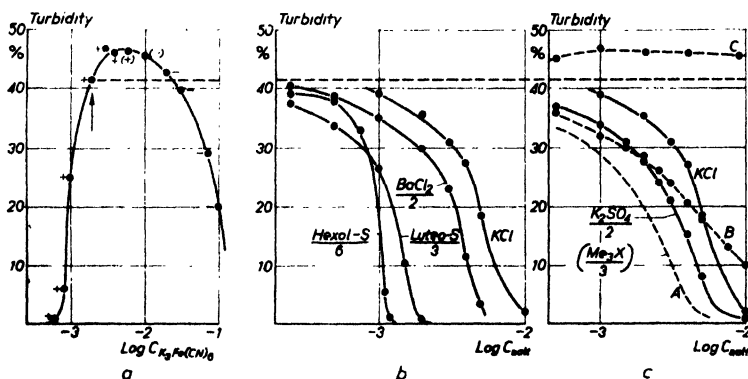
Fig. 40. Coacervation of 0.04°₀ ichthyocoll with K₃Fe(CN)₆ at pH = 2.92 (A) and suppression of this coacervation by salts (see text).
The reversal of charge point in A lies at the top of the turbidity curve.
The arrow here indicates the working point for the influence, investigated in B and C, of added salts. The curves A, B and C are discussed a little further on in the text (p. 391).

point was practically proportional to the concentration of the gum arabic. This was caused by the very strong binding of the hexol ions through which the corresponding "true concentrations" are very small (of the order of 10^{-6} N).

Now this is no longer the case with the combination positive ichthyocoll + $K_3Fe(CN)_6$. Here one finds only a slight increase of the required "concentrations" of the $K_3Fe(CN)_6$ on increasing the ichthyocoll concentration. This is caused by the much less intense binding of the $Fe(CN)_6'''$, through which the "true concentrations" are appreciably greater and of the same order of magnitude as the gross concentrations. The true reversal of charge concentration here lies round about the order of 5 m. eq. per 1 (log $C = 0.70$ —3).

This point of difference is not attributable to the fact that here we have an example of the variant colloid cation + micro anion, in the other case an example of the variant colloid anion ⋰ micro cation but to the choice of the micro ions. Roughly

speaking one can say that this difference is to be attributed to the valency of the micro ion (6-valent compared with 3-valent), compare p. 265, Fig. 4).

The true reversal of charge concentrations may with both variants lie both at very low and at higher concentrations. In general they increase with decreasing valency of the active micro ion. We shall return to this in § 3g and § 3h (p. 400, 404).

c. Anomalies in the double valency rule and their significance

In the above given example of positive ichthyocoll + $Fe(CN)_6'''$ the normal valency rule of the anions has been reduced to only two terms, since $K_3Fe(CN)_6$ and also $K_4Fe(CN)_6$, that is to say, 1 — 3 and 1 — 4 contain anions which themselves contract powerful complex relations with the ichthyocol cation, as a result of which dicomplex coacervation is produced.

These salts therefore are differently situated than K_2SO_4 and KCl, the anions of which bring about no or very weak complex relations at least in the concentrations used for the suppression so that the suppressive action of both cation and of anion is quite prominent with K_2SO_4 and KCl.

$K_3Fe(CN)_6$ nevertheless also suppresses the coacervation which at smaller concentrations it has itself brought about but for this auto-suppression relatively very great concentrations are necessary. Compare Fig. 40a and b from which it is seen that only above 100 m. eq. per 1 (log C = 0.00 —1) suppression by $K_3Fe(CN)_6$ is as powerful as that with KCl already below 10 m. eq. per 1 (log C = 0.00 —2).

At the chosen working point of 2 m. eq. per 1. $K_3Fe(CN)_6$ extra $K_3Fe(CN)_6$ added even produces at first an increase of the concervervation (since the working point in the left graph still lies before the top of the curve).

From this one calculates the curve C in the extreme right graph for extra added $K_3Fe(CN)_6$ (Fig. 40c). This curve will thus fall again at much higher concentrations, but this will take place appreciably only between 30 and 300 m. eq. per 1.

If one wishes to include $K_3Fe(CN)_6$ also in the valency rule of the anions in the suppressive action of the salts, one would have to write:

$$1 - 2 \; \rangle \; 1 - 1 \; \rangle \; \ldots\ldots\ldots\ldots \; 1 - 3 \; (K_3Fe(CN)_6)$$

in which the long dotted line indicates that $1 - 3$ $(K_3Fe(CN)_6)$ suppresses the coacervation only at very much higher concentrations than the other two.

Now these considerations are of importance for the explanations of the occurrence of abnormal valency sequences which are frequently encountered in the suppression of dicomplex coacervation by salts.

The following example is typical of this. The positive ichthyocoll sol is not brought to coacervation by potassium methane trisulphonate $(K_3CH(SO_3)_3)$. One may therefore expect that coacervation with $K_3Fe(CN)_6$ is suppressed by $K_3CH(SO_3)_3$. The dotted line B in the right graph of Fig. 40 reproduces the result. This curve takes an abnormal course, because it does not lie as one would actually expect (curve A) according to the normal valency rule of the anions but on the contrary intersects the curves for K_2SO_4 and KCl.

If we compare the order of the salts at the level of 35% turbidity and a number of lower turbidity levels, we then find first the normal valency rule of the anions (a), but afterwards all kinds of abnormal orders (b-e) until finally an order is produced which is quite comparable with what was discussed above (f). See survey below.

(a) $1 - 3 \rangle 1 - 2 \rangle 1 - 1$ (d) $1 - 2 \rangle 1 - 3 = 1 - 1$

(b) $1 - 3 = 1 - 2 \rangle 1 - 1$ (e) $1 - 2 \rangle 1 - 1 \rangle 1 - 3$

(c) $1 - 2 \rangle 1 - 3 \rangle 1 - 1$ (f) $1 - 2 \rangle 1 - 1 \rangle \ldots . 1 - 3$

The explanation of this gradual shift of $1 - 3$ to the right in the series of symbols is now clear. The $CH(SO_3)_3'''$ ion is similarly to the $Fe(CN)_6'''$ ion also able to contract complex relations with the ichthyocoll cation. These complex relations however are appreciably weaker and are still hardly felt in the region of very small concentrations (below 1 m. eq. per l). Here the $K_3CH(SO_3)_3$ has a stronger suppressive action than K_2SO_4. On increase of the concentration the complex relations increase in intensity and cause the above given displacement in the series of symbols towards the right. The maximum reachable intensity of the complex relations between $CH(SO_3)_3'''$ and ichthyocoll cation is however much weaker than in the similar combination with $Fe(CN)_6'''$ since here a dicomplex coacervation does not take place.

Similar anomalies in the order of the terms of the double valency rule can naturally also occur in the variant micro cation + colloid anion but they then naturally manifest themselves in the rule of the cations. Even a glance at Fig. 37 shows that when we include the auto-suppression of the hexol nitrate in the valency rule of the cations we encounter similar features here: $3 - 1 \rangle 2 - 1 \rangle 1 - 1 \rangle \ldots . 6 - 1$.

Also rhodochromic chloride (type $5 - 1$) and $Pt(en)_3(NO_3)_4$ (type $4 - 1$) give dicomplex coacervation of the gum arabic sol so that one cannot expect that these salts will settle themselves in the normal cation series in the suppression of dicomplex coacervation with $6 - 1$. This is therefore here restricted to only the first three terms.

We shall see later (p. 396, § 3f) that on replacing the aqueous medium by water-alcohol or water-acetone mixtures similar term displacements can occur in this series of three terms (p. 398, Fig. 43) as we have discussed above for the terms of the anion series with the positive ichthyocoll sol.

In general such displacements for a salt, which itself causes no dicomplex coacervation or flocculation, therefore indicate that the cation or anion in question already produces fairly powerful complex relations, although these are not yet powerful enough to cause dicomplex coacervation or flocculation. Therefore in this case *dicomplex sols* of the type colloid anion + micro cation or colloid cation + micro anion are produced.

d. Further examples of dicomplex coacervation or flocculation, colloid ion + oppositely charged micro ion. Role of the ion valency and of the colloid equivalent weight

Orienting investigations have shown that a very great number of cases in which hydrophilic sols are coacervated or flocculated by salts, belong to this type of dicomplex coacervation or flocculation[1].

[1] H. G. BUNGENBERG DE JONG and collaborators, see publications in Bioch. Z. 235 (1931) 185; 248 (1932) 115; 248 (1932) 131; 248 (1932) 309; 248 (1932) 335; 254 (1932) 15; 254 (1932) 35; 257 (1933) 62; 259 (1933) 436; 259 (1933) 442; 260 (1933) 161; 262 (1933) 390; 263 (1933) 33; further *Rec. trav. chim.*, 53 (1934) 607; 53 (1934) 622; 53 (1934) 737; 53 (1934) 747; 54 (1935) 1; 54 (1935) 17. In these publications the coacervation or flocculation in question were still denoted by the terms „autocomplex coacervation or flocculation" now dropped and the experimental data are discussed from the standpoint now also obsolete that the kinetic units of the sols have the nature of a dispersed phase.

Negative colloids (macromolecular colloids: acidoids and negative proteins, as well as association colloids e.g. phosphatides) very frequently give coacervation or flocculation with hexol nitrate (see for example p. 244, Fig. 6) in which as appears from their properties are due to the production of complex relations between the colloid anion and the hexol cation. However not only the hexol cation is able to do this but also other cations. In general one can say that the chance of a dicomplex coacervation or flocculation of this type increases with increase of the charge of the cation. This does not prevent lower valency cations being able to do this also in special cases. But not only inorganic monatomic or complex cations are able to do this but also organic cations although they are only monovalent, for example, basic dyes[1] or alkaloid cations[2] (for example p. 235, Fig. 4; p. 466, Fig. 34 and pag. 467, Fig. 35).

In all these examples coacervation or flocculation ranges occur depending on the concentration with reversal of charge points in these ranges. The coacervates exhibit typical deformation and disintegration phenomena in an electric field. Furthermore the double valency rule and the continuous valency rule are also always encountered. It does not however hold any longer for these cations that the true reversal of charge concentrations are always extremely small. We have therefore in general the case schematized in Fig. 36c instead of that in Fig. 36b (p. 386).

Positively charged proteins give similar dicomplex systems of the type *colloid cation + micro anion* with suitably chosen anions[1].

The chance of their occurrence is again great with polyvalent inorganic anions (for example $Fe(CN)_6''''$), with large organic anions of complicated structure (for example acid dyes) and also with less complicated anions such as picrate and sulphosalicylate.

Here also there are again the same indications of the dicomplex nature (coacervation or flocculation ranges with a reversal of charge point lying within them, suppression by added neutral salts, valency rules, disintegration phenomena in a d.c. electric field)[1].

If we now go into further detail, the valency of the oppositely charged cation and the equivalent weight of the colloid stand out as striking factors in the occurrence or otherwise of dicomplex coacervation or flocculation with the inorganic micro ions. In the following survey the behaviour with regard to some six salts[3] is indicated for a number of negative colloids by the position of a vertical line. The salts to the left of this line give dicomplex coacervation or flocculation, the salts to the right of it are not able to do this.

	Reciprocal hexol number
Na pectate	
6 — 1, 5 — 1, 4 — 1, 3 — 1, 2 — 1 | 1 — 1	203
Na semen lini mucilage	
6 — 1, 5 — 1, 4 — 1, 3 — 1 | 2 — 1, 1 — 1	563
Na arabinate	
6 — 1, 5 — 1, 4 — 1 | 3 — 1, 2 — 1, 1 — 1	1068
Amylum solubile	
| 6 — 1, 5 — 1, 4 — 1, 3 — 1, 2 — 1, 1 — 1	26000

[1] H. G. BUNGENBERG DE JONG and J. LENS, *Bioch. Z.* 254 (1932) 15.

[2] H. G. BUNGENBERG DE JONG and C. VAN DER MEER, *Proc. Koninkl. Nederland. Akad. Wetenschap.*, Amsterdam, 45 (1942) 593.

[3] 6—1 = hexol nitrate; 5—1 = rhodochromic chloride; 4—1 = $Pt(en)_3(NO_3)_4$; 3—1 = $Co(NH_3)_6Cl_3$; 2—1 = $CaCl_2$; 1—1 = NaCl.

Alongside the colloids are given the reciprocal hexol numbers (p. 270, as a measure of the colloid equivalent weight) determined on the colloid preparations used.

We thus see that a cation with lower valency can still be used for the production of the dicomplex coacervation or flocculation the smaller the equivalent weight of the colloid, that is to say the more ionised groups are present per gram of colloid or more correctly when we are dealing with a "linear" colloid the greater the linear occupation density for ionised groups along the macromolecule.

For the internal state of the coacervate (or the flocculation) we again have to bear in mind similarity with the variant colloid cation + colloid anion with salt formation, although of probably short life in the case of thoroughly liquid coacervates (in this case therefore between the ionised groups of the macromolecule and the dicomplex forming ion).

As far as the significance of the ion valency is concerned one can among other things, consider also that polyvalent ions, can form bridges between ionised groups of two or more neighbouring colloid ions. Nevertheless polyvalency of the micro ion is not a *conditio sine qua non* since cases are also known in which monovalent ions (e.g. dye cations, picrate anions) cause complex coacervation or flocculation (see further p. 404, § 3h).

e. Specific factors in the dicomplex coacervation or flocculation colloid ion + oppositely charged microion. Equivalent flocculation

The above indicated regularity is however already broken when we do not restrict ourselves to a small number of biocolloids but pass in review the whole of the material available to us regarding flocculability with the six chosen salts.

The reader should reread in Chapter IX, § 1 d and § 1 e (p. 269-273) bearing in mind that the term flocculability used there firstly makes no distinction between coacervation and flocculation, secondly that all the "flocculations" mentioned there belong to the dicomplex type colloid anion + micro cation.

It then appears that beside the colloid equivalent weight another specific factor is present which depends on the nature of the ionised group of the colloid anion. One must distinguish between phosphate, carboxyl and sulphate colloids and for the same colloid equivalent weight the flocculability also decreases in this order.

However the representation given in § 3 d (p. 392) appears in another way also to be too simple. It is not true that the valency of an ion is the only factor which is of importance for dicomplex coacervation or flocculation. We know numerous examples in which ions of the same valency nevertheless behave differently. We have already discussed a single example in § 3c (p. 391) where we saw that the positive ichthyocoll sol does give dicomplex coacervation with $K_3Fe(CN)_6$ but not however with $K_3CH(SO_3)_3$ although this is also a salt of the type 1—3.

Na nucleate furnishes a further example. As may be expected from its fairly low equivalent weight, it flocculates with cations of low valency (Fig. 41):

$$6—1, \quad 5—1, \quad 4—1, \quad 3—1, \quad 2—1 \mid 1—1.$$
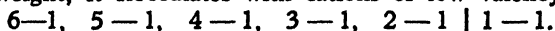
The symbol 2 — 1 comprises $MgCl_2$, $CaCl_2$, $SrCl_2$ and $BaCl_2$ in this case[1].

If now one investigates adequately low concentrated nucleate sols, it is seen that only $CaCl_2$ and $BaCl_2$ give flocculation while the sols remain clear with $MgCl_2$ and $SrCl_2$ (not published)[2].

---

[1] H. G. BUNGENBERG DE JONG and F. A. MENALDA, *Bioch. Z.* 257 (1933) 62.

[2] Very diluted nucleate sols may remain clear even with $CaCl_2$ (as in Ch. IX, § 6e, p. 330).

From this and many other examples it appears that not only the valency of a micro ion is of importance but also other specific factors play a role. Since now the the reversal of charge phenomena play a central role in dicomplex systems and one in fact encounters in the "ion spectrum" of Na nucleate (p. 283, Fig. 13c) for the alkaline earth metals the order for reversal of charge concentration increasing from left to right:

$$Ba - Ca - Sr - Mg,$$

it is clear that the study of the "ion spectra" must be considered to be of the greatest importance for answering the question as to the specific behaviours of the ions. We shall return to this later in § 3g and 3h, p. 400-405).

The example put forward above of dicomplex flocculation of nucleate with cations, gives us the opportunity to discuss yet another point, namely the so-called *equivalent flocculation* which is often encountered in dicomplex flocculations. This phenomenon is present here for hexol nitrate and $Co(NH_3)_6Cl_3$. Compare the coincidence of the rising branches of the turbidity curves in Fig. 41. The name indicates that mutually equivalent amounts of electrolyte are necessary. Another example concerns the flocculation of Na thymus nucleate [2] in which besides the above

Fig. 41. Flocculation of Na yeast nucleate with some salts[1].
Ordinates: turbidity.
Abscissae: logarithms of the salt concentration in eq. per 1.
The rising branches of the turbidity curves on the left for hexol nitrate (6 — 1) and $Co(NH_3)_6Cl_3$ (3 — 1) almost coincide.

mentioned salts $Pt(en)_3(NO_3)_4$ (type 4 — 1) was included in the investigation. Here equivalent flocculation was observed for all three 6 — 1 = 4 — 1 = 3 — 1. For an explanation we have to consider that the concentrations at which this occurs are really quite fictitious.

We discussed this already for the coacervation limit of gum arabic with hexol nitrate in a (see p. 386, Fig. 36b, curve A). Practically all added hexol ions are here bound to the gum arabic and the "true" coacervation concentration is very much lower (of the order of 10^{-6} N).

Even so the true flocculation concentrations in Fig. 41 for 6—1 and 3—1 are very much lower than the concentrations indicated. These latter represent for 6 — 1 and 3 — 1 therefore practically the whole of the cations bound to the nucleate. With a 2 × concentrated sol the rising turbidity branch will thus be shifted over 0.3 (that is log 2) because for just producing the flocculation twice as many ions must now also be bound. Similarly this curve will be displaced over 0.3 to the left if the sol concentration is twice as small as the original one.

Two salts, although their "true" flocculation concentrations can be very different, will nevertheless just flocculate at practically the same gross concentrations when

---

[1] One would be inclined to expect that the Ba curve in the figure would be just the curve rising highest. However it should be borne in mind that the order of the alkaline earth cations is a „transition series" (p. 289) and as a result Ba and Ca are not entirely comparable. With Ca the polarising influence predominates, with Ba rather the size of the cation.

[2] H. G. BUNGENBERG DE JONG and ONG SIAN GWAN, *Kolloid Beih.*, 31 (1930) 89.

these latter are great compared with the two mutually still different true concentrations. This is apparently the case for 6 — 1, 4 — 1 and 3 — 1 in the above mentioned examples. At the equivalent flocculation the (true) concentrations are thus not equivalent but the amounts of oppositely charged ions bound to the colloid are equivalent.

Equivalent or almost equivalent action of salts with polyvalent ions is also frequently to be observed already in the lowering of the relative viscosity [1] (see for example p. 224 and 225, Fig. 34b and 35).

f. Dicomplex coacervation or flocculation supported by alcohol or acetone

We again choose the gum arabic sol for the discussion of the point mentioned in the title. We remind the reader that for this sol in an aqueous medium we meet for the six salts mentioned in § 3d (p. 392)

$$6 — 1, \quad 5 — 1, \quad 4 — 1 \mid 3 — 1, \quad 2 — 1, \quad 1 — 1$$

that is to say, with the salts to the left of the vertical line (hexol nitrate, rhodochromic chloride, $Pt(en)_3(NO_3)_4$) dicomplex coacervation occurs while with the salts to the right ($Co(NH_3)_6Cl_3$, $CaCl_2$, $NaCl$) coacervation or flocculation does not take place with any concentration.

Since now however we saw in § 3d that other colloids can indeed be brought by 3 — 1 or by 3 — 1 and 2 — 1 to dicomplex coacervation of flocculation, it is obvious to suppose that with gum arabic these salts do indeed bring about true complex relations but that these — in view of the relatively high colloid equivalent weight — are not considerable enough to be able to force the separation as coacervate or flocculi against the forces which otherwise hold the arabinate in solution.

When one starts to replace the medium consisting exclusively of water by media which from a general viewpoint are less suitable solvents, it may be expected that dicomplex coacervation nonetheless occurs with the salts which stand to the right of the vertical line in the series of symbols given above. This is indeed the case in alcohol-water or acetone-water media [2]. One may compare Fig. 42, which relates to 0.087% sols and in which the attached numbers relate to the constant acetone concentration for each curve (namely cc. acetone to 25 cc final volume).

We see that in media containing 0—18 vol. % acetone $Co(NH_3)_6Cl_3$ does not bring about turbidity at any single concentration but at 20% acetone ("5") there is for the first time a range of concentrations in which a weak turbidity is produced On further increase of the acetone concentration this turbidity becomes stronger and the range of salt concentration in which this occurs also expands more and more.

We also establish the same for $CaCl_2$, only here the minimum acetone concentration required is greater here ("7" not yet, but already yes at "7.5") and further the maxima of the turbidity curve, at least for those curves in which the top has not yet become very flat, lie at considerably greater concentrations.

Finally the same also occurs with NaCl, in which again still greater acetone concentrations are nevessary and the maxima appear at still greater salt concentrations.

Microscopic investigation shows that flocculi or lumps separate when the turbidity is strong enough but coacervate drops when this is weak.

[1] H. G. BUNGENBERG DE JONG and J. LENS, Bioch. Z., 235 (1931) 174.
[2] H. G. BUNGENBERG DE JONG and K. C. WINKLER, Bioch. Z., 248 (1932) 115; 259 (1933) 436.

It is further of interest that reversal of charge from negative to positive takes place in these turbidity ranges[1].

With sufficiently large coacervate drops the deformation and disintegration phenomena in an electric field could be observed in their typical forms.

Finally it appears that salts can also have a suppressive action according to the "double valency rule", in which naturally the abbreviations of the series of terms of the cations already discussed in § 3c (p. 391) and anomalous sequences of these terms occur. Compare Fig. 43 which relates to the flocculation with 2 m. eq. per l $Co(NH_3)_6Cl_3$ throughout in various acetone concentrations. Here the suppressive actions of the two first terms of the cation series (2 — 1 and 1 — 1) and of the anion series (1 — 2 and 1 — 1) have been investigated.
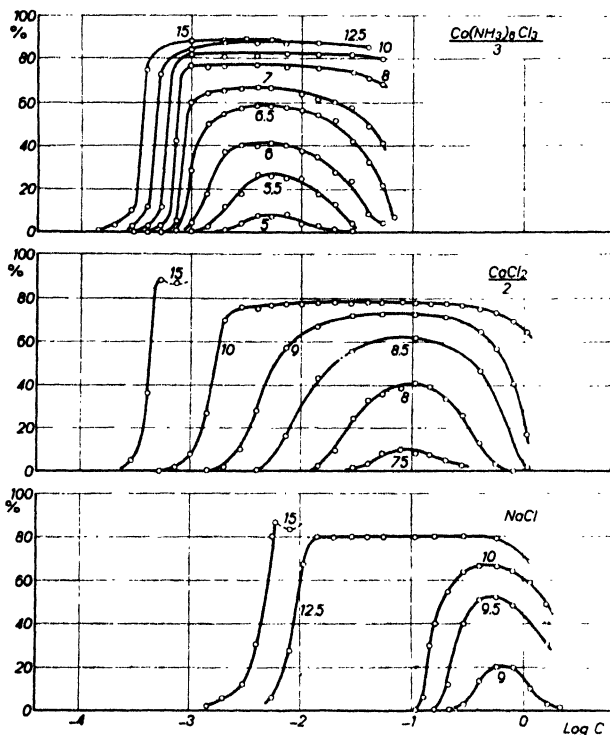
Fig. 42. Flocculation (or coacervation) of 0.087% gum arabic sol with $Co(NH_3)_6Cl_3$, $CaCl_2$ and NaCl in water-acetone mixtures of varying composition.
Ordinates: turbidity.
Abscissae: log. of the salt concentrations (in eq. per l). The added numbers indicate how many cc of acetone are present per 25 cc of final mixture.
At 15 cc only the rising branch of the curve on the left can be measured reproducibly, the further course (not indicated) is irregular on account of strong floccule formation.

At smaller acetone concentration (x = 6) the order of the terms is still normal for both:

$$2 — 1 \rangle 1 — 1 \quad \text{and} \quad 1 — 2 \rangle 1 — 1,$$

although the $CaCl_2$ curve already approaches the NaCl curve more closely at its bottom end than higher up.

At x = 7 intersection of these curves already takes place, although at this acetone concentration $CaCl_2$ itself still just does not cause flocculation.

---

[1] If the turbidity is slight the reversal of charge point lies on the right-hand falling branch of the curve, with greater turbidity near the distinct turbidity maximum or in the flat maximum tract at still higher acetone concentrations.

On further increase of the acetone concentration this abnormal course of the $CaCl_2$ curve becomes ever more conspicuous. This is therefore caused by the complex relations, which Ca itself creates, interfering with the suppressive action. To illustrate this further the turbidity curve of $CaCl_2$ if it was present alone has been drawn dotted in the graphs for $x = 7.5$, 8 and 9, and similarly that for NaCl in $x = 9$.

While therefore abnormalities occur in the series of terms of the cations, the order of the terms in the anion series:

$$(1 - 2 > 1 - 1)$$

is maintained throughout.
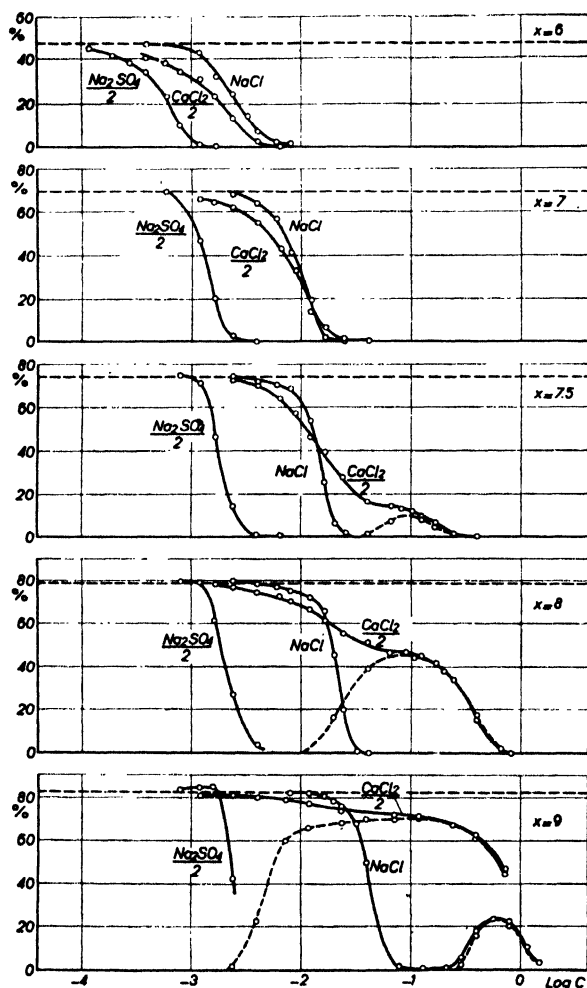
Fig. 43. Suppression of the flocculation (coacervation) of gum arabic with 2 m. eq. per l. $Co(NH_3)_6Cl_3$ by added salts at increasing acetone concentration. x: the number of cc of acetone present per 25 cc of the final volume.
Ordinates: turbidity; the turbidity caused by 2 m. eq. per l $Co(NH_3)_6Cl_3$ in the absence of salts: is indicated by the horizontal dotted lines.
Abscissae: the logs of the salt concentration (in eq. per l). At $x = 7.5$ or higher $CaCl_2$ itself already gives flocculation (coacervation) (dotted curve), similarly NaCl at $x = 9$ (dotted curve).

Fig. 44 gives the suppression of the $CaCl_2$ flocculation (4 m. eq. per l.) in 40 vol. % acetone by a number of salts.

Naturally the series of terms of the cations is here restricted to the first member $(1 - 1)$. Since at this acetone concentration NaCl already gives flocculation (the dotted curve) the course of the NaCl curve is understandable. At smaller concentrations it has at first the usual suppressive action on the Ca flocculation but at higher concentrations the Na ion causes sufficiently strong complex relations. The curve does not fall to zero but rises again and from then on follows the dotted curve: the NaCl turbidity.

The order of the terms of the anion series is perfectly normal 1—3 ⟩ 1—2 ⟩ 1—1 and the curves for 1 — 3 and 1 — 2 completely suppress the turbidity. This happens because this suppression occurs at smaller concentrations than NaCl and consequently

Fig. 44. Suppression of the CaCl₂ flocculation in 40% acetone by added salts. CaCl₂ concentration constant = 4 m. eq. per l.
The turbidity corresponding to this is given by the dotted horizontal line.
Ordinates: turbidity.
Abscissae: logarithms of the salt concentration (in eq. per l.). The dotted curve gives the turbidity already occurring with NaCl itself (in the absence of CaCl₂).
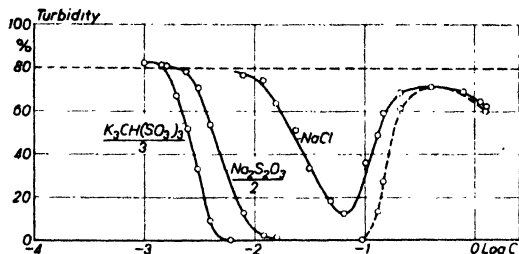
the extra complex relations which the K or Na ion will bring about at higher concentrations, are here not yet noticeable.

Summarising we can say that the coacervations or flocculations which $Co(NH_3)_6Cl_3$, $CaCl_2$ and $NaCl$[1] bring about are of a *dicomplex* nature. They are of the same kind as 6 — 1, 5 — 1 and 4 — 1 cause in aqueous medium.

Through a gradual increase of the acetone concentration, from the original series of terms with the vertical line between 4 — 1 and 3 — 1 there are produced series in which the line each time is placed more towards the right:

    25 vol % acetone
6 — 1, 5 — 1, 4 — 1, 3 — 1 | 2 — 1, 1 — 1
    32 vol % acetone
6 — 1, 5 — 1, 4 — 1, 3 — 1, 2 — 1 | 1 — 1
    38 vol % acetone
6 — 1, 5 — 1, 4 — 1, 3 — 1, 2 — 1, 1 — 1 |
that is to say, increase of the acetone or alcohol concentration formally acts in the same way as decrease of the equivalent weight of the colloid anion (see p. 393).

A deficit of ionised groups on the macromolecule can thus be compensated by changing the solvent in the direction in which the colloid is finally completely insoluble. Here one can in the first place think of a decrease of the solution forces which play their part between the various hydrophilic groups on the macromolecule and the solvent.

Closer investigation shows however that for a not so small part the acetone or the alcohol is also active in another direction namely in a promotion of the complex relations themselves.

This is already more or less apparent in Fig. 42 (p. 397) from the extention of the flocculation ranges with CaCl₂ or NaCl on increase of the acetone concentration which happens especially towards the left. However a direct indication of this is the shift of the reversal of charge concentrations towards lower concentrations on increase of the acetone or alcohol concentration which occurs not only with gum arabic but also with all the acidoids investigated[2]. Compare e.g. on p. 403, Fig. 47b with Fig. 47a.

---

[1] The complex nature is also seen from the presence of disintegration phenomena of the coacervat drops.
[2] Many examples are given in the publications in the *Bioch. Z.* cited in note on p. 392.

g. Mutual comparison of the inorganic cations as regards the dicomplex coacervation or flocculation of gum arabic

In § 3f we have considered the dicomplex coacervation or flocculation with a more or less arbitrarily chosen representative of each of the salt types 6 — 1, 5 — 1, 4 — 1, 3 — 1, 2 — 1, 1 — 1 in aqueous medium and in alcohol-water or acetone-water mixtures. As was already shown in § 3e (p. 394) however from a few examples, specific differences
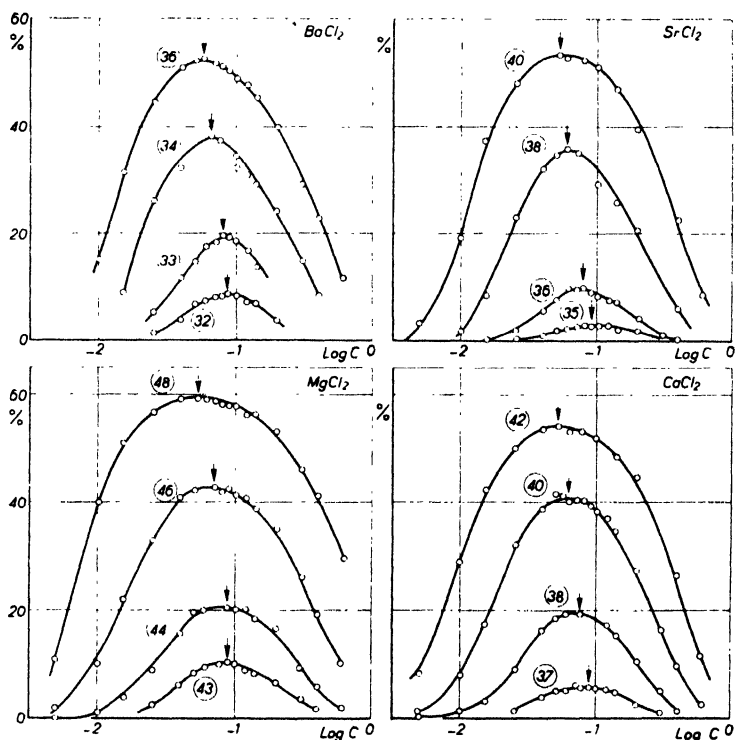
Fig. 45.   Turbidity ranges with the chlorides of Mg, Ca, Sr and Ba in gum arabic sols containing alcohol.
Ordinates: turbidity.
Abscissae: logarithms of the salt concentrations in eq. per 1.
Arrows: salt concentrations at which the turbidity maximum lies.
Numbers beside the curves give the constant alcohol concentrations (in vol % 96% alcohol).

also occur between cations of the same valency and we shall now have the opportunity of going further into the question with the help of a more extensive collection of data. The test object is again gum arabic. Experiment shows that hexol nitrate (6—1), rhodochromic chloride (5 — 1), $Pt(en)_3(NO_3)_4$ (4 — 1) give dicomplex coacervation in aqueous medium; further that $Th(NO_3)_4$ (4 — 1) $AlCl_3$ (3 — 1) give flocculation

in aqueous medium, on the other hand the sol remains clear at all concentrations with $Co(NH_3)_6Cl_3$, $Ce(NO_3)_3$ and $La(NO_3)_3$ all three of the type 3 — 1. The latter is the case as well with the nitrates or chlorides of all the remaining divalent cations (Mg, Ca, Sr, Ba, Co, Ni, Mn, Cd, Zn, UO_2, Cu, Pb) and monovalent cations (Li, Na, K, NH_4, Ag).

The flocculations with $Th(NO_3)_4$ and $AlCl_3$ form a separate category of complex flocculation because the flocculations once produced are no longer completely reversible by added indifferent salt.

Since here the strong tendency to hydrolysis also plays a part[1] we presumably are dealing with a flocculation of the negative arabinate with positive thorium or aluminium hydroxide. These flocculations will not occupy us any further here but we shall turn to the remaining 3, 2 and 1-valent cations which do not coacervate or flocculate in aqueous medium.

Now it holds for each of these salts that on gradually increasing the alcohol or acetone concentration coacervation or flocculation ranges occur, as we saw to be the case for Co $(NH_3)_6Cl_3$, $CaCl_2$ and NaCl in Fig. 42 (p. 397). In these ranges turbidity curves with a still well defined maximum occur in the first place but at higher alcohol or acetone concentration these maxima become too flat for their exact position still to be determinable accurately. Fig. 45 gives four such sets of curves from which the existence of specific differences between cations of the same valency already appears from the alcohol concentrations written by the curves (Ba $\rangle$ Sr $\rangle$ Ca $\rangle$ Mg). With all the above mentioned ions similar sets of curves have been determined with gum arabic and from them the salt concentrations at which the maximum lies for a given alcohol concentration[2]. The results of this investigation are collected in Fig. 46.

Each of the curves drawn in it gives the run of the electrolyte concentration for optimum turbidity as a function of the alcohol concentration. Each of these curves ends below at a critical alcohol concentration, at which turbidity no longer occurs. The curves are again difficult to follow further upwards than is given since there the turbidity curve gets too flat a top.

The slope from bottom right to top left shows that the attachment of the cations to the gum arabic is strengthened by increasing alcohol concentration. When now we look at the order of the curves from left to right, that is to say at constant alcohol concentrations, then one notices first of all the influence of the valency of the cation. The curves for the 3-valent cations lie far to the left and begin at the smallest alcohol concentration (Al even at 0%, that is to say, in aqueous medium itself, see above however).

The curves for the 2-valent cations lie more in the middle, while far to the right are encountered those for monovalent cations. We record this regularity as

$$3 — 1 \rangle 2 — 1 \rangle 1 — 1 \qquad\qquad (a)$$

and mean by this that the complex relations between arabinate anion and micro cation are the more powerful the higher the valency of the cation. This is also underlined by the order of the alcohol concentrations, required for each group just to cause the appearance of the turbidity, decreasing in this order.

---

[1] H. G. BUNGENBERG DE JONG, W. A. L. DEKKER and K. C. WINKLER, *Rec. trav. chim.*, 53 (1934) 607.
[2] H. G. BUNGENBERG DE JONG and R. STOOP, *Kolloid chem. Beih.*, 42 (1935) 96.

Nevertheless a single glance at the figure shows that striking differences occur between cations of the same valency. First of all we may mention the more powerful action of the ions of the B sub-groups compared with those of the A sub-groups:

$$UO_2, Pb, Cu, Cd, Zn \text{ etc. } \rangle Ba, Sr, Ca, Mg$$

$$\text{and } Ag \rangle K, Na, Li \tag{b}$$
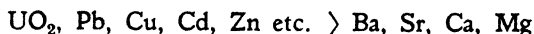
In addition there are also specific differences which are connected with the volume of the ions:

With the trivalent ions we meet the order

$$Al \rangle Ce \cdot \rangle La \rangle Co(NH_3)_6 \tag{c}$$
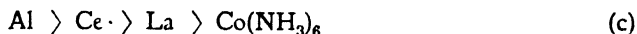
in which the ion volume increases from left to right. With the divalent cations of the A subgroups the following order holds on the other hand:

$$Ba \rangle Sr \rangle Ca \rangle Mg \tag{d}$$

In this case as with the monovalent cations

$$K \rangle Na \rangle Li \tag{e}$$

the cations act precisely the less powerfully the smaller their volume.
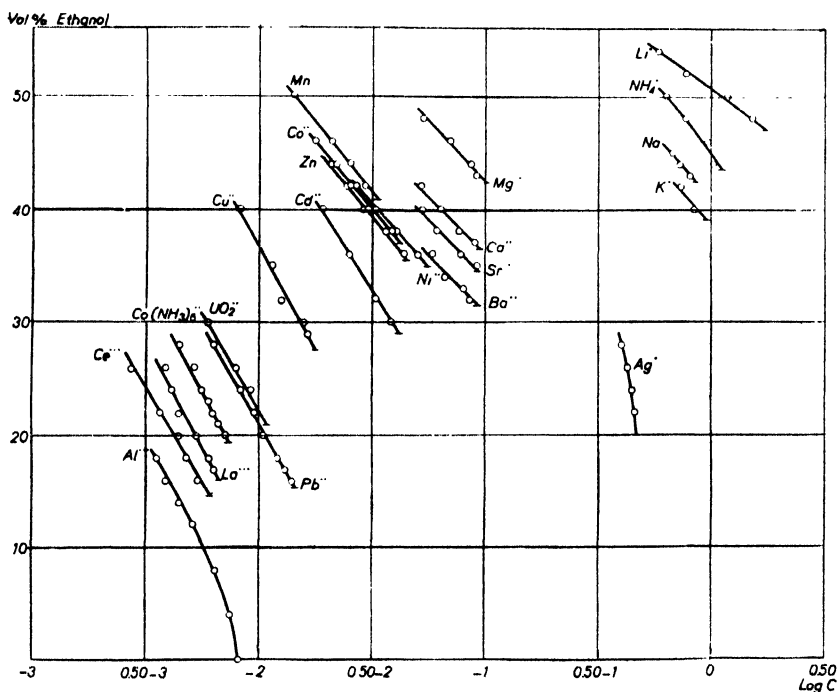
Fig. 46. Logarithms of the salt concentrations (in eq. per l) at various alcohol concentrations, at which the turbidity maximum lies for the flocculation of gum arabic with salt + alcohol. (see text).

These regularities (*a* to *e*) can be explained by assuming:

(1) that salt formation, that is to say, a binding of the cations to the ionised carboxyl groups of the gum arabic is fundamental for the flocculation.

(2) that the carboxyl group is more strongly polarisable than the water molecule.

The sequences mentioned under (a), (b) and (c) are indeed just those which one would expect when an ionised group more strongly polarisable than water is postulated.

Actually one would also expect the sequences mentioned under (d) and (e) to be reversed. That this however did not occur is an indication that the polarisability of the COO' group of the arabinate is only but little stronger than that of water. One is concerned here in fact with large and relatively large ions, the polarising action of which is the smallest. Sequences occur here which on neglecting polarisation effects are precisely the ones which may be expected as a consequence of the decrease of the hydration of the cations with increasing ion volume.

In view of the close connection between the occurrence of coacervation or flocculation ranges and reversal of charge phenomena it was to be expected that the regularities summarised above would also appear in the concentrations at which the flocculi or coacervate drops are just reversed in charge. For technical reasons this investigation
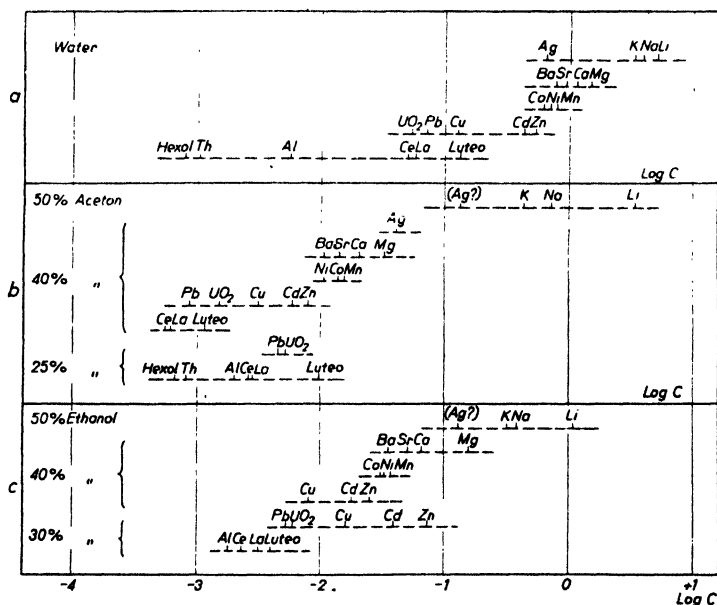
Fig. 47. Cation sequences for the reversal of charge of Na arabinate in water (a) or in 25, 40, and 50% acetone (b) and for the flocculation maxima of gum arabic in 30, 40 and 50% alcohol (c). The concentrations in c were taken from Fig. 46 (intersections of the curves with horizontal lines at 30, 40 or 50% alcohol).

a. Since, apart from hexol, Th and Al, the arabinate does not flocculate with any of the other salts, the reversal of charge was here determined electrophoretically using suspended SiO_2 particles, which surround themselves with a complete arabinate film.

b. The reversal of charge is here determined on the flocculi or coacervate drops themselves.

was carried out however in acetone-water mixtures, as appears permissible, since acetone plays quite the same role as alcohol (p. 396, § 3f). The results for certain constant acetone concentrations (25, 40 and 50 vol %) are collected in Fig. 47b (p. 403) while for comparison the salt concentrations are plotted in a similar way in Fig. 47c at which according to Fig. 46 (p. 402) the maximum flocculation is situated at certain constant alcohol concentrations (30, 40 and 50 vol %). On comparison of Fig. 47b and 47c we see that with at most a single exception (Ni, Co) the same orders of the cations occur.

If now we return to the exclusively aqueous medium, then it is obvious to suppose that the cations which do not bring about coacervation or flocculation, nevertheless will bring about complex relations at suitable concentrations, but these will be too weak to produce coacervation or flocculation. Evidence, for this is, among other things, the occurrence of minima with these salts in the relative viscosity. See p. 217, Chapter VII, § 9 and p. 220. Fig. 31.

But a direct proof would be also lie in the detection of reversal of charge in the clear sols. Here the formation of colloid films on suspended quartz particles (p. 277) comes to our assistance and it now appears that one can indeed obtain reversal of charge in exclusively aqueous medium with all salts.

The cation sequences obtained, see Fig. 47a, now again appear to be essentially the same as in b and c. Furthermore it is important that the reversal of charge concentrations are throughout considerably lower for the same cation in media containing acetone (b) than in water (a). It appears once more from this that acetone or alcohol strenghthens the attachment of the cations.

h. Significance of the "reversal of charge spectra"

The good agreement present in § 3g between the sequences of the cations in dicomplex coacervation or flocculation on the one hand and the sequences of these ions as regards the reversal of charge concentrations on the other hand, have led to a great significance being attached to these latter as general information regarding the complex relations between a given colloid and ions of the opposite sign.

A propos of this these reversal of charge sequences have been determined more extensively with all kinds of colloids and the *"reversal of charge spectra"* so obtained can serve as the basis for the interpretation of all kinds of phenomena which depend on interrelations between colloids and ions.

The results have already been reproduced *in extenso* in Chapter IX (inorganic ions p. 276, § 2 and 3; organic ions p. 300, § 4) to which reference may be made.

As regards the dicomplex coacervation, colloid ion + oppositely charged ion, it follows from it that we no longer consider the factor provisionally put in the foreground in § 3d (p. 392), the ionic valency, as being quite so important but rather the place which an ion occupies in the reversal of charge spectrum. By this we can manage to include the organic cations or anions in a general discussion.

We saw indeed that these ions, although all monovalent for example, can still differ considerably in reversal of charge concentration (compare e.g. the fatty acid anions in Fig. 31, p. 307, or the alkaloid cations in Fig. 23, p. 301).

In general one may then consider that with a given colloid the chance of dicomplex coacervation or flocculation increases the lower the reversal of charge concentration of the ion in question.

We have in fact always to bear in mind that the ions — at any rate the ion with the same electrical sign as the colloid — of the salt, with which the dicomplex coacervation or flocculation is brought about, has a suppressive action. When thus the (true; p. 276) reversal of charge concentration is small the complex relations can come strongly into being; if the reversal of charge concentration is greater or is great then the complex relations at the reversal of charge point turn out to be smaller or no longer reach the value which is at least necessary for coacervation or flocculation.

Besides that the other factor put forward in § 3d (p. 392) retains its full weight. The smaller the equivalent weight the more will the colloid tend to dicomplex coacervation or flocculation under otherwise comparable conditions (e.g. comparing colloids with the same ionized groups[1]). Both factors in fact play an important part in the dicomplex coacervation or flocculation of phosphate, carboxyl and sulphate colloids with alkaloid ions[2].

i. Dicomplex coacervation of phosphatides with cations

Phosphatides can under certain conditions[3] give coacervation or flocculation with cations which to judge by all kinds of properties belong to the dicomplex type. We shall not go into details on this point[4] but merely mention a few peculiarities. The cation sequences are typically those of the phosphate colloids (see for more detail p. 280, Ch. IX, § 2c; p. 289 and 295): e. g.

$$Ca \ \rangle \ Mg \ \rangle \ Sr \ \rangle \ Ba \ \ldots \ldots \ \rangle Li \ \rangle \ Na \ \rangle \ K$$

As was to be expected the water content of the coacervates increases from left to right in these series[5] (each time compared at the concentrations of optimum coacervation). The reversal of charge concentrations of the divalent cations (for example Ca) are relatively small compared with other colloids, those of Na and K on the contrary are fairly large.

A striking peculiarity of these coacervates is that when once they have been produced they are not or difficulty reversible[6]. Extra added salts do indeed bring about reversible changes in the water content in the direction which may be expected. For example the water content of a coacervate produced with Ca is increased by NaCl. This coacervate naturally decreases in water content (recognizable for example by vacuolation occurring in the coacervate drops) on the addition of a salt with a cation which occupies a place further to the left in the reversal of charge spectrum than Ca. These coacerv-

---

[1] See already p. 296, Ch. IX, § 2m, from which it seems presumable that the above rule will not hold in comparing sulphate colloids using a strongly polarizing cation.

[2] H. G. BUNGENBERG DE JONG and C. VAN DER MEER, *Proc. Koninkl. Nederland. Akad. Wetenschap.*, Amsterdam, 45 (1942) 593. The results suggest further that the strength of the individual salt bond ionized group-large (i.e. non polarizing) cation still depends on the composition of the former: with equal charge density a phosphate colloid has greater flocculability than a sulphate colloid, and the latter than a carboxyl colloid.

[3] Support either by substances soluble in water such as alcohol, resorcinol etc., or by substances insoluble in water such as triolein and cholesterol. These latter sensitizers must be already incorporated in the kinetic units during the preparation of the sol. See note 1 on p. 274 and further literature quoted in note 4.

[4] See H. G. BUNGENBERG DE JONG and R. F. WESTERKAMP, *Bioch. Z.*, 248, (1932) 131, 309, 335.

[5] H. G. BUNGENBERG DE JONG and G. G. P. SAUBERT, *Proc. Koninkl. Nederland. Akad. Wetenschap.*, Amsterdam, 45 (1942) 40.

[6] This is the more remarkable because the dicomplex variant protein cation ·|· negative phosphatide is reversible. The tricomplex flocculations in which phosphatides take part as amphoions can also be thoroughly reversible (see p. 415, § 6a and p. 429, § 6f).

ates thus act as ion exchangers, whereby the changes produced in the intensity of the complex relations give rise to corresponding changes in the water content.

A further peculiarity of these phosphatide coacervates is that the disintegration phenomena are of an abnormal type, although this is again with positive drops the mirror image of that of negative ones (see Ch. XI § 6b p. 479).

Finally we mention as a last peculiarity that these coacervates can also change their water content under the influence of all sorts of organic non-electrolytes[1]. In this respect they are similar to coacervates of oleates which are obtained by salting out (KCl)[2]. This sensitivity has nothing to do with the dicomplex nature but is closely connected with the presence of long hydrocarbon chains of the esterified fatty acids in the phosphatide molecule (oleate coacervates are further treated in Chapter XIV).

The abnormal behaviour of the phosphatides is closely connected with the association character of these colloids. Their negative charge is due to impurities of an acid nature (probably phosphatidic acids) which are tenaciously attached to the actual phosphatide molecules by London — v. d. Waals forces. The relative amount of them compared with the actual phosphatide molecules (amphoion!) determines their colloid equivalent weight, the magnitude of their reversal of charge concentrations with cations and the position of the isoelectric point[3] (see p. 295).

Even a very small admixture of these acid substances (hardly or not at all detectable in the N: P ratio) results in a large shift of the isoelectric point, as is understandable when one considers that, for example with egg lecithin, both the acid and the basic group in the phosphatide molecule has a very large dissociation constant (see p. 191-192).

j. Dicomplex coacervation colloid ion + micro ion reduced to its simplest form

The usual way, in which a dicomplex coacervation of this type is brought about, is that a suitably chosen salt is added to a "hydrophilic" sol. Since the colloid is itself also a salt, consisting of colloid ion + micro ion this dicomplex coacervation is according to the Phase Theory an "demixing" in an at least *quaternary system*. (H_2O + salt I + salt II, in which the salts I and II have no common ion; see already p. 367 and 387). Of the four ions present there are only two essential complex partners, for example, the underlined ones in the following combinations.

Na arabinate + hexol nitrate, or clupein chloride + $K_3CH(SO_3)_3$.

Clupein is a protein with a low equivalent weight and in agreement with this it coacervates with $K_4Fe(CN)_6$, $K_3CH(SO_4)_3$, K_2SO_4 and even by addition of NaCl.

In this connection the following data taken from KOSSEL[4] are of importance for us: clupein sulphate is readily soluble in hot water but separates out from the

[1] H. G. BUNGENBERG DE JONG and G. G. P. SAUBERT, Protoplasma, 28 (1937) 329.

[2] H. G. BUNGENBERG DE JONG, H. L. BOOIJ and G. G. P. SAUBERT, *Protoplasma*, 28 (1937) 543.

[3] The soya bean phosphatide insoluble in alcohol has a very low I. E. P. and also a very low colloid equivalent weight, with which is connected the fact that even in aqueous medium (without addition of auxiliary substances, see note 3, p. 405) complex flocculation occurs with 3 and 2-valent cations. According to the N:P ratio = 2 : 1 one can here expect a considerable admixture of phosphatidic acid.

H. G. BUNGENBERG DE JONG, G. VERBERG and R. F. WESTERKAMP, *Kolloid. Z.*, 71 (1935) 194.

[4] A. KOSSEL, *The protamines and histones*, Longmans, Green and Co, London, New York, Toronto, 1928, see p. 30, 31.

solution as a clear colourless oil. At room temperature this oil contains about 50% water. The overlying solution contains about 1.3 — 1.6% clupein sulphate. This description fits perfectly with the phenomenon that we have called coacervation: the sol separates into a colloid-rich and a colloid-poor liquid layer.

An investigation of the properties of dilute clupein sulphate sols gives evidence that clear complex relations are present in them between the clupein colloid cation and the sulphate anion[1] (see in more detail Ch. VII, § 12, p. 228).

The "demixing" of concentrated clupein sulphate sols on cooling is therefore a dicomplex coacervation colloid cation + micro anion in its simplest form, namely as "demixing" in the binary system H_2O + salt.

We read further in KOSSEL that clupein chloride is soluble in water but is separated as an oil by added NaCl. Obviously the complex relations here between colloid cation and the monovalent chlorine ion are still too weak for coacervation in the binary system clupein chloride + H_2O and without further information one cannot determine whether the "demixing" by added NaCl is based on increase of these complex relations, or on an ordinary salting-out process.

In this connection it is perhaps useful to point out again (see already p. 226) the difference in principle between salting out and transgression of solubility by complex formation, which can greatly resemble each other phenomenologically.

Salting out occurs in general at high electrolyte concentrations (e.g. 1N and higher). The charge density of the colloids plays no part in it. The effectiveness of the ions is governed by the well known lyotropic series.

On the other hand we find complex flocculation or coacervation rather at relatively low electrolyte concentrations. They are the more pronounced the greater the charge density of the colloid. The effectiveness of the ions is in the main determined by their valency over which lyotropic influences are superposed. We sometimes find here that the sequence of the ions is just the opposite of that in the normal lyotropic series (for example of the monovalent anions, with the positive proteins, compare also p. 299, Fig. 22) and we have in this a clear indication of the complex character of the flocculation.

§ 4. DICOMPLEX SYSTEMS, III. THE VARIANT MICRO CATION — MICRO ANION IN CONNECTION WITH THE THEORY OF COMPLEX COACERVATION

a. "Demixing" in mixtures of aqueous salt solutions

Although the theory of complex coacervation in the narrower sense (p. 370, § 2o) was originally developed on the basis of the assumption that sols are in principle two phase systems, good reasons were present for assuming that a theory which is based on the representation of the biocolloid sols as true solutions would contain a greater element of truth.

This latter representation had been based on the fact that it is not only possible to replace one of the two colloids by a suitably chosen micro ion ("crystalloid" ion) of the same sign of charge while retaining the typical properties of complex coacervation in the narrower sense (§ 3) but also and especially on the fact that we have

[1] H. G. BUNGENBERG DE JONG, W. A. L. DEKKER and P. VAN DER LINDE, Rec. trav. chim., 54 (1935) 1.

succeeded in finding examples in which both colloids could be replaced by organic or inorganic micro ions[1].

The "demixing" of hexol nitrate + $K_3Co(CN)_6$ has been investigated more extensively whereby numerous points of complete similarity with complex coacervation were established[2].

On mixing 10 m. eq. per 1 solutions of the two salts "demixing" occurs in a range of mixing proportions. The microscopic "demixed" drops readily fuse with each other into larger ones, they flow out over the glass surface (wetting) and eventually a very dark coloured viscous liquid layer is obtained by sedimentation.

The "demixed" drops carry an electrophoretic charge on their surface, they are positively charged with excess of hexol nitrate, negatively with excess of $K_3Co(CN)_6$. The reversal of charge point lies close to the equivalent mixing proportion. Though with difficulty, disintegration phenomena of the drops in an electric field could be observed. The "demixing" is suppressed by indifferent salts according to the double valency rule:

$$1 - 2 \; \rangle \; 1 - 1 \text{ and } 3 - 1 \; \rangle \; 2 - 1 \; \rangle \; 1 - 1.$$

The electrophoretic velocity of the drops is influenced by added salts according to the continuous valency rule:

(relative   (relative
positivation) $3 - 1 \ldots 2 - 1 \ldots 1 - 1 \ldots 1 - 2 \ldots 1 - 3$ negativation)

Compared with complex coacervation of colloid cation + colloid anion the behaviour is thus completely similar, although no colloid at all is present.

There are only two points of difference:

1. The "demixing" is here associated with certain minimum concentrations (about 2 to 3 m. eq. per l) of the two salts solutions. In the complex combination gelatin + gum arabic coacervation takes place at pH 3.5 on mixing equal volumes of 0.001% sols. The limit with this combination lies below 8 μ eq. per l[3]. This point of difference is however of a quantitative nature and not a matter of principle.

2. The "demixed" phase is metastable and crystals are formed after standing longer. With the complex coacervation gelatin + gum arabic (and others) no indications of metastability of the coacervate are ever observed.

[1] H. R. KRUYT and H. G. BUNGENBERG DE JONG, Proc. Koninkl. Nederland. Akad. Wetenschap., Amsterdam, 38 (1935) 714 discussed the extension of the theory of complex coacervation to micro-ionic systems, based on the following examples:
$SrCl_2$ — ammonium molybdate.
$BaCl_2$ — ammonium molybdate.
$Cd(NO_3)_2$ — Na tartrate.
$Cd(NO_3)_2$ — Na benzoate.
$Cd(NO_3)_2$ — Na succinate.
Luteo cobalti chloride — antimonyl potassium tartrate.
Hexol nitrate — Na amylsulphonate (iso).
Hexol nitrate — $K_3Co(CN)_6$.
H. G. BUNGENBERG DE JONG and L. TEUNISSEN-VAN ZIJP, Proc. Koninkl. Nederland. Akad. Wetenschap., Amsterdam, 39 (1936) 1103, in a qualitative investigation found 34 more cases of „demixing" in aqueous salt solutions.
[2] H. G. BUNGENBERG DE JONG and K. C. WINKLER, Z. anorg. allgem. Chem., 232 (1937) 119.
[3] On assuming an equivalent weight of 1200 for arabinic acid the figure 8 μ eq. per l follows for the concentration of 0.001% sol. At pH 3.5 the (apparent) equivalent weight is however greater so that the concentration expressed in equivalent becomes still smaller.

b. Stable and metastable demixing in binary systems salt + water

The problem of complex coacervation is considerably simplified by the existence of similar demixing phenomena in micro ionic solutions. It follows then that specific colloid chemical considerations based, for example, on the macromolecular structure of the colloids are not really essential. The variants of dicomplex coacervation discussed in the previous paragraphs, together with the variant micro cation + micro anion can all be formulated as a double decomposition

$$AB + CD \rightarrow AD\, n.aq \quad BC$$
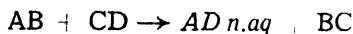
in which the difficultly soluble salt AD separates in an amorphous form, that is to say, as an "demixed" liquid containing water:

I.   Gelatin chloride + Ca arabinate →
$\qquad\qquad$ *gelatin arabinate n. aq.* $CaCl_2$
II. Gelatin chloride + Na picrate →
$\qquad\qquad$ *gelatin picrate n. aq.* $NaCl$.
III. Hexol nitrate + Ca arabinate →
$\qquad\qquad$ *hexol arabinate n. aq.* $Ca(NO_3)_2$
IV. Hexol nitrate + potassium cobalticyanide →
$\qquad\qquad$ *hexol cobalticyanide n. aq.* + KNO_3.
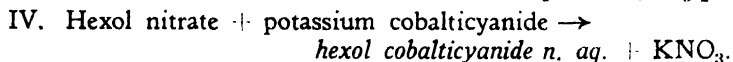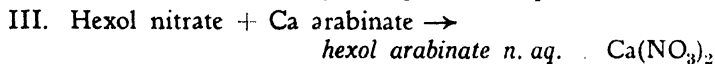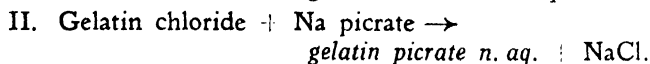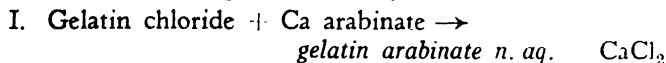
Although these formulations are analogous to the usual double decomposition for the case in which a hydrated salt is precipitated, for example,

$$CaCl_2 + Na_2SO_4 \rightarrow CaSO_4 . 2aq. + 2\, NaCl$$

there is nevertheless a fundamental difference.

In the crystalline hydrated salt the salt ions and the water are arranged in a lattice, in the amorphous coacervates (I—III and the microionic analogue IV) there is no question of a strictly ordered arrangement over larger distances. The fact that coacervates vary continuously in composition under the influence of all sorts of variables (mixing proportion of the colloids, pH, added salts etc.) can be attributed to this.

In this connection we must also remark that the above mentioned formulations are not correct, not even at equivalent ratios of the electrolytes to the left of the arrow, because the neutral salt BC which remains in solution after the double decomposition, still distributed over both coexisting liquids. Thus the formulations are only approximately correct at very low concentrations. When the mixing proportions are not equivalent, the (colloid) salt, which is present in excess, is in addition distributed over both coexisting liquids.

The existence of this great variability in the composition of the two liquid phases is not surprising according to the Phase Theory because the system H_2O + salt AB + salt CD already forms a quaternary system (see p. 366, § 2m).

If now it is correct that only two of the four ions present are essential for complex coacervation (A and D) then an "demixing" must be able to occur even in the binary system H_2O + AD. We have already met a single example (H_2O + clupein sulphate, see p. 407). It is therefore important for the theory of complex coacervation to investigate if a similar "demixing" is realisable in the case in which both A and D are micro ions. Evidence for their existence (Sr molybdate, luteocobalti thiosulphate)

had been obtained previously[1] but on account of the readiness to decompose of these salts they are unsuitable for quantitative investigation. A number of crystalline salts derived from the nitrogen base β diethylaminoethyl p-aminobenzoate (the hydrochloride of which is well known under the name of procaine hydrochloride)[2] were found to be suitable for the work.

They can be obtained by double decomposition of the hydrochloride with an alkali salt of the anion in question. In this case demixing first occurs, but at room temperature the salt-rich liquid layer is again metastable and the anhydrous salt crystallises out. This salt can be further purified by crystallisation.
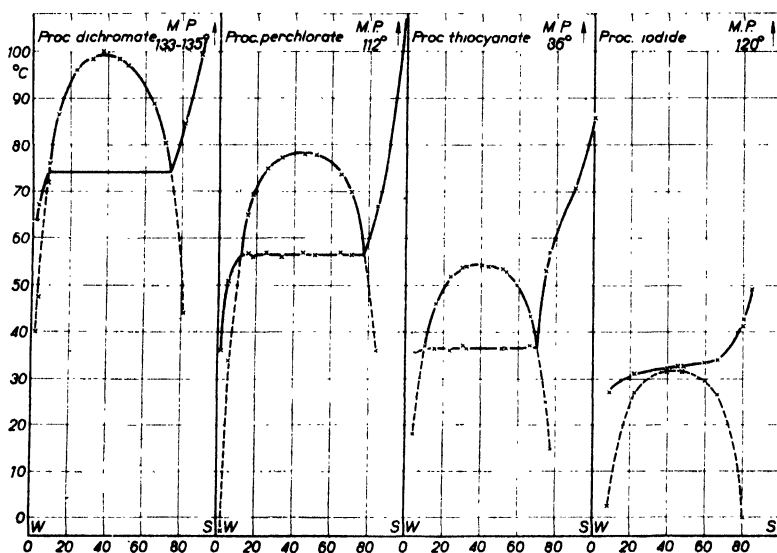
Fig. 48. t — x diagrams for the binary systems of four procaine salts with water.

With 4 such salts the temperature — composition diagram (t — x) has been investigated for the binary systems water + salt by means of the so-called "thermal analysis" method (see Fig. 48). Weighed amounts of the anhydrous salt and water are slowly heated in a sealed ampoule or cooled and the temperature noted at which changes occur in the number or kind of the phases. In the diagrams (Fig. 48) the ordinate axis gives the temperatures (°C), the abscissa axis the composition in percentage by weight of the salt in the total system.

The significance of the curves in the diagram is well known. Let us discuss, for example, the case of procaine perchlorate. Starting from a 50% mixture at room temperature we have saturated solution + anhydrous crystals.

[1] H. R. KRUYT and H. G. BUNGENBERG DE JONG, Proc. Koninkl. Nederland. Akad. Wetenschap., Amsterdam, 38 (1935) 714.

[2] L. W. J. HOLLEMAN and H. G. BUNGENBERG DE JONG, Rec. trav. chim., 59 (1940) 1055.

On gradually raising the temperature this state continues to exist until 56°, at which a three-phase equilibrium occurs: salt-poor liquid + salt-rich liquid + anhydrous crystals. On further rise of temperature the crystals disappear and we have only two coexisting liquids: salt-poor + salt-rich solution.

Finally above 78° these make way for one homogeneous liquid system. On cooling the changes occur in the reversed order, thus first "demixing" again at 78°, etc. But here it is fairly easy to keep the unmixed state at temperatures below the three-phase temperature at which really stable equilibria exist only between anhydrous crystals and salturated solution.

With the aid of thermal analysis the limits of the metastable "demixing" region can indeed be established. These are represented by the dotted branches of the curves, which are the continuation of the (full line) "demixing" curve in the temperature region (56°—78°) in which the "demixing" is stable. The other two branches of the curves represent the equilibria saturated solution — anhydrous crystals.

We see that the critical "demixing" temperature (top of the "demixing" curve) decreases in the order Cr_2O_7'' — ClO_4' — CNS' — I'. The distance (in degrees Centigrade) between critical "demixing" temperature and the temperature of the three-phase equilibrium also decreases in the same order and even disappears completely in the case of the iodide, in which the top of the "demixing" curve coincides approximately with the solubility curve of the crystals. The "demixing" is metastable throughout in this case.

"Demixing" could not be established in the binary systems bromide—water, nitrate—water or chloride—water. It is possible that these also have "demixing" regions since under certain circumstances, namely in the presence of a salt with the same anion (for example procaine bromide — K Br) at high concentration an "demixing" can occur. The metastable "demixing" regions in question would then have to be sought in the binary system at temperatures far below the crystallisation temperature.

If we summarise the above statements we can state that:

1. "Demixing" may occur even in binary systems H_2O + micro ionic salt.

2. This "demixing" needs not always be metastable but can in some examples be stable in a certain temperature range.

3. The critical temperature of demixtion for the case of the given organic cation depends on the anion. It decreases in the sequences: I ⟩ Br, Cl and CNS ⟩ NO_3.

We need not again repeat the significance of point 1. for the theory of complex coacervation (the three variants treated in § 2 and 3). As far as point 2. is concerned we note that the only apparently essential point of difference of complex coacervation with the "demixing" micro cation — micro anion (cited on p. 408 under 2). disappears.

And finally in point 3. the same anion sequences appear as we met in the reversal of charge spectrum of positive proteins (protein cations, see p. 299).

Summarising we may therefore consider these and similar "demixings" in binary systems H_2O + salt as complex coacervation reduced to its simplest form: beside water only the two essential ions are present; these in the case of procaine salts are only monovalent and do not even possess macromolecular structure.

c. The essence of complex relations

Complex coacervates sometimes possess a relatively high water content and the diagrams of Fig. 48 (p. 410) invite one to make a comparison with the water content of the "demixed" salt-rich liquid. Naturally this latter water content depends on the temperature chosen but for a comparison with the water content of the complex coacervate gelatin (positive) + gum arabic (negative) it is indicated to choose a temperature sufficiently far below the critical solution point, for example 20°, so that the composition does not change much per degree fall in temperature. This because at temperatures at which the water content of the gelatin — gum arabic coacervate is known, one is far removed from a possible critical solution point[1] and the water content at those temperatures is almost constant (p. 341).

The water-poorest complex coacervate (pH = 3.5, mixing proportion 50%) only contains about 18% colloids, thus more than 80% water. From the diagrams of Fig. 48 we read off for the procaine salt-rich layer at 20° on the other hand a salt content of about 75—85% and this layer therefore contains only 15—25% water.

The considerations of § 2o (p. 370) are useful for explaining this enormous difference. In this section we saw that as a consequence of the macromolecular structure of the colloids a large amount of water is present as water of occlusion in the statistical molecular clews.

We there omitted any further discussion on the question of what more detailed representations one must form of the interaction of the ionised groups of both signs, but only made use of the idea that these oppositely charged groups approach each other very closely.

Here there are still two possibilities:

1. They actually lie adjacent to each other.

2. They approach each other up to a short distance but are throughout separated by water of hydration.

At first sight 2. appears to be preferable because then the liquid nature of the coacervate and the free relative displacement of the colloids in an electric field (disintegration phenomena) are explained without further argument.

One can also put forward in favour of 2. that the salt-rich layers in the diagrams of Fig. 48 still contain 15—25% water.

This point of view 2. resembles to the original idea: electrostatic attraction versus hydrative repulsion (p. 370), but now restricted to hydration of the ionogenic groups.

There is strong evidence in favour of the point of view 1. in the variants colloid anion + micro cation or colloid cation + micro anion in the specific ion sequences for reversal of charge or for coacervation or flocculation. We remind the reader for example of the sequences Cs < Rb < K < Na < Li which occur with sulphate colloids and carboxyl colloids (p. 289). Here polarisation phenomena are still in the background and here the largest ion, that is to say, the least hydrated ion, is most suitable for reversal of charge or coacervation. This points strongly therefore to a direct contact between cation and ionised group of the colloid.

With the phosphate colloids the sequence is a transition series. For the largest

[1] Coacervation still occurs at 100° C, for example.

ions which cannot polarise, the order is the same: Cs $\langle$ Rb $\langle$ K but for the smaller ions just the opposite Li $\langle$ Na $\langle$ K (p. 289-290).

Of the last three it is just the smallest ion, thus the most heavily hydrated ion, which is in the most favourable position. Nevertheless it does not owe this most favourable position to its stronger hydration but to the larger polarising action of the anhydrous ion. The phosphate group is considerably more polarisable than water. The Li ion actually wants to get rid of its water of hydration — although this costs energy — because a much larger polarisation energy is set free on the attachment of the anhydrous ion to the phosphate group.

The sequence Li $\rangle$ Na $\rangle$ K arises in this way. This and other sequences point to a direct contact of cation and ionised group of the colloid. Similarly the sequence I — Br — Cl with protein cations (p. 299) points in the same direction. Since here polarisation of the anions no longer takes place it is again the largest ion, i. e. the least hydrated anion, which is best attached and is the most favourable for coacervation. In this connection it is interesting that with the procaine halides (large organic cation!) also the iodide is again in the most favourable condition as regards "demixing". It is therefore improbable that the 15—25% water in the salt-rich liquids of Fig. 48 is situated exclusively and permanently between the ionised group of the procaine ion and the anion as we postulated in point of view 2 (p. 412).

Thus when one accepts point of view 1. for the variants colloid cation + micro anion and colloid anion + micro cation there is, on account of the very great similarity, no reason for doubting that it also holds for the variant colloid cation + colloid anion (i. e. for the complex coacervation in the narrower sense).

When we weigh the two points of view against one another, preference is given to the first, that is to say, to a direct contact of the oppositely ionised groups and (or) ions. However the liquid nature of the coacervates (the gelatin — gum arabic coacervate is a Newtonian liquid) and the relative displaceability of the colloids in an electric field (disintegration) demand a fuller explanation. For these reasons the direct contacts postulated in point of view 1. can hardly be of a static kind, that is to say, be permanent. We must rather conceive these contacts as dynamic, that is to say, they are of relatively short duration and every now and again let go to make afresh other contacts once more.

For the variant micro cation + micro anion we can conceive the salt-rich layer in the same way. This layer can more or less be conceived as an ion lattice extremely strongly distorted by intruding water, whereby actually a "molten" salt containing water has been produced. The disturbances increase more and more in extent on rise of the temperature until finally the two layers become completely miscible.

§ 5. UNICOMPLEX SYSTEMS

a. Unicomplex sols

Since a protein is not really uncharged at the isoelectric point but the same numbers of positively and of negatively charged ionised groups are present here side by side one has thus to reckon with the fact that at or in the immediate neighbourhood of the isoelectric point complex relations are present between these groups. As is known two cases present themselves with regard to this, the protein remains in solution at the isoelectric point: *isostable proteins*, or the protein separates out of the solution: *isolabile proteins*.

Sols of the first group, for example, of gelatin can be regarded as *unicomplex sols*, if at least there is evidence to be found in their behaviour which points to this. Such evidence is present in the influence already discussed in Chapter VII, § 9, p. 217 of small concentrations of neutral salts on the relative viscosity of the iso-electric gelatin sol. These, namely, increase the relative viscosity and at smaller concentrations the greater the valency of the ion. One can therefore conceive this as a suppression of the complex relations.

An indirect indication is also that, when one adds just enough alcohol till the isoelectric gelatin sol becomes weakly opalescent, this opalescence disappears on addition of a little indifferent salt. Further the turbidity of cooled isoelectric sols which disappears with small salt concentrations.

As regards the question why the gelatin sol is isostable, one must answer that the complex relations between the molecules themselves are not large enough to surpass the solubility. The loose macromolecular structure possibly plays a part in this, as a result of which a large part of the interrelations between positively and negatively charged groups takes place intra molecularly and only a small fraction remains available for intermolecular inter-relations.

b. Unicomplex flocculation

The isolabile proteins do flocculate, coacervate or crystallise at their isoelectric point. For a large group of these proteins, which have received the name of globulins, it is characteristic that they only separate out at their isoelectric point, when the medium contains no, or sufficiently little, indifferent salt. *The protein separated out at the I.E.P. has the property of dissolving in dilute salt solutions.* These properties give us a direct proof for the idea that the separation of the globulins at the I.E.P. is based on the existence of pronounced intermolecular complex relations. Indeed the property of indifferent salts of suppressing complex relations has already been met with very frequently in this chapter. It is then to be expected
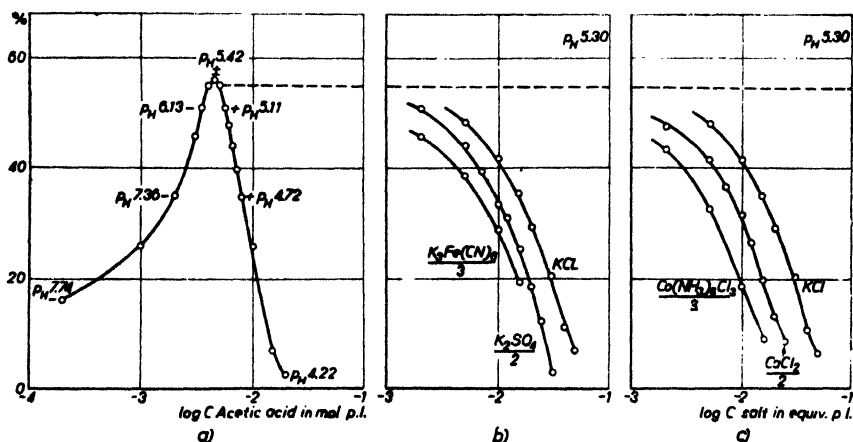
Fig. 49. Flocculation of serum globulin from dilute horse serum in the neighbourhood of the isoelectric point (A) and suppressive action of indifferent salts (B and C) close by the maximum of the curve in A. Ordinates: turbidity.

that the double valency rule (p. 350) will have to appear also in the dissolving action of salts on globulin flocculations. Indeed it has been found in an investigation of the isoelectric flocculation of serum globulin[1] that at constant concentration (m. eq. per l) the solvent action increases from right to left in the following series (see Fig. 49).

$$3 - 1 \rangle 2 - 1 \rangle 1 - 1 \qquad 1 - 3 \rangle 1 - 2 \rangle 1 - 1.$$

For the interpretation of individual differences between ions of the same valency (lyotropic influences) in the "salting in" of globulins the considerations on ion sequences as treated in § 2 s, p. 376 (c. f. also p. 429, § 6 f.) may also be helpful.

As, however, the globulins belong to the class of corpuscular proteins, which have been excluded from treatment in this book, we will not enter further into these problems (for extra complications see p. 261-262).

§ 6. TRICOMPLEX SYSTEMS

a. Tricomplex flocculation. Working hypothesis

In the continuation of the investigation on complex flocculation or complex coacervation a type was found which deviates completely from the types known up to now[2].

It is characteristic of this type that three essential components must be present simultaneously.

A clear sol of purified egg lecithin (association colloid) does not flocculate immediately with $La(NO_3)_3$[3]. Na-arabinate sol also gives no flocculation or coacervation with $La(NO_3)_3$. The sol mixture of egg lecithin and Na arabinate similarly is clear. If however one adds a solution of $La(NO_3)_3$ to a suitably chosen mixture of egg lecithin and Na arabinate sol, then flocculation immediately occurs in a definite range of concentrations. This flocculation completely bears the character of a complex flocculation. It becomes quickly suppressed on the addition of indifferent salts, (e.g. $Co(NH_3)_6Cl_3$, $CaCl_2$, $NaCl$, K_2SO_4, $K_3CH(SO_3)_3$), in which the double valency rule occurs:

$$3 - 1 \rangle 2 - 1 \rangle 1 - 1 \qquad \text{and } 1 - 3 \rangle 1 - 2 \rangle 1 - 1.$$

Further there occurs in the flocculation range reversal of charge of the flocculi from negative to positive at a certain $La(NO_3)_3$ concentration.

And as a third characteristic of the complex nature of the flocculation the continuous valency rule is encountered in the influence of added indifferent salts on the electrophoretic velocity of the flocculi:

(relative   (relative
positivation) $1 - 3 \ldots 1 - 2 \ldots 1 - 1 \ldots 2 - 1 \ldots 3 - 1$ negativation)

The example given here does not stand alone. One can replace each of the three; the egg lecithin, the salt and the acidoid, by suitably chosen others while retaining the typical complex flocculation (which thus only occurs if all three

---

[1] H. G. BUNGENBERG DE JONG, W. A. L. DEKKER and K. C. WINKLER, *Rec. trav. chim.*, 53 (1934) 607.

[2] H. G. BUNGENBERG DE JONG and G. G. P. SAUBERT, *Bioch Z.*, 288 (1936) 1.

[3] Opalescence only occurs after a long time.

components are present simultaneously, on the other hand is absent when they are combined two by two).

The following survey gives a few such combinations, which can be considered as produced from the first by replacement

egg lecithin   — La(NO₃)₃ — Na arabinate

egg lecithin   — Ce(NO₃)₃ — Na arabinate

egg lecithin   — La(NO₃)₃ — K chondroitin sulphate

isoelectric gelatin — La(NO₃)₃ — Na arabinate

isoelectric gelatin — Zn(NO₃)₂ — K chondroitin sulphate.

Since in these combinations $K^·$ or $Na^·$ is always present together with NO_3' ions, it is important to notice that neither egg lecithin nor isoelect. gelatin nor arabinate nor chondroitin sulphate flocculates with KNO_3 of $NaNO_3$ and that these salts bring about flocculation just as little in mixtures of these sols.

It is already clear from this that the $K^·$, $Na^·$ or NO_3' ions can play no part (or on the contrary only a weakening role) in causing complex flocculation. For this there thus remain the constituents underlined each time in each combination.

The isoelectric gelatin is present as an amphoion and obviously the egg lecithin here also plays its part as an amphoion. In agreement with this is also the fact that one can in fact obtain similar flocculations with soya bean phosphatide, but these are less intense and the reversibility of the flocculation by added neutral salts leaves something to be desired.

We already discussed previously (see p. 270, Table 2; p. 274 and 295) that egg lecithin, acting as a negative association colloid, has a very high equivalent weight, the soya bean phosphatide a much lower one. That means that the egg lecithin is much less mixed with acid constituents (for example phosphatidic acid) than the soya bean phosphatide and thus approximates much better to the theoretically pure phosphatide, which must consist exclusively of amphoions.

We arrive with this at the pronouncement that for the production of the flocculations discussed here the simultaneous presence of complex relations between a *colloid amphoion*, a suitably chosen *micro cation* and a suitably chosen *colloid anion* is essential.

It has now further appeared that this colloid anion can also be replaced by suitably chosen micro anions, while retaining the typical complex flocculation[1].

The following list gives a small selection of the many examples found:

egg lecithin   — CaCl₂   — (NH₄)₆Mo₇O₂₄

egg lecithin   — CaCl₂   — K₂HgI₄

egg lecithin   — Cd(NO₃)₂ — KCNS

isoelectric gelatin — CaCl₂   — (NH₄)₆ Mo₇O₂₄

isoelectric gelatin — CaCl₂   — K₂HgI₄

isoelectric gelatin — Cd(NO₃)₂ — KCNS.

[1] H. G. Bungenberg de Jong and G. G. P. Saubert, *Bioch. Z.*, 288 (1936) 13.

We can summarise the points discussed so far by saying that in these flocculations complex relations are present between three essential types of ion,

amphion — cation — anion,

in which in the known examples the anion is either a micro anion or a macromolecular colloid anion, the cation is a micro cation and the amphoion is either a macromolecular amphoion (protein) or an association of micro amphoions (phosphatide).

From the characteristic that these three ion types simultaneously play off complex relations against each other, these systems have been called *tricomplex systems*[1].

Now the working hypothesis presents itself more or less spontaneously that — at least in non complicated cases[2] — the cation puts itself in complex relation to the negatively charged ionised group of the amphoion, the anion on the other hand to the positively charged ionised group of the amphoion.

Compare Fig. 50 in which the uni-, di- and tricomplex systems have been symbolised as simply as possible by the schematic figures I, II and III. In them cations and
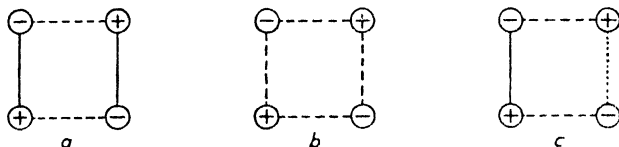
Fig. 50.   Schemes for the unicomplex (a), dicomplex (b) and tricomplex (c) systems (see text).

anions have been taken as monovalent and possible multipolar ions have been simplified to dipolar ions. The complex relations present between them have been indicated by dashes.

Although the dicomplex systems can be symbolised by a figure in which only one cation and one anion are depicted, their number is doubled to illustrate the mutual connection with the two other schemes (in which a total of four charges occurs).

The complex relations, which occur in the scheme for the unicomplex systems, are mutually equal, as also those occurring in the scheme for the dicomplex systems.

This is not the case with the tricomplex systems where the three complex relations indicated differ in intensity from each other. This is not in fact completely manifest in this scheme because the complex relations between cation and anion are merely indicated in a different way, namely by dots, those between each of them and the charges of the zwitter ion again by dashes.

The purpose of this is discussed further below.

---

[1] For a fuller discussion of this choice and a survey of the theoretically possible variants see § 1 p. 335.

[2] It is quite possible that some heavy metal cations, possibly also the $UO_2^{..}$ ion, attach themselves to a non-ionogenic group of the protein molecule and this combination of amphoion and cation, since it now carries more positive than negative charges, begins to behave as a colloid cation with respect to the anion and then really forms a dicomplex system with it.

With very simple cations, for example, Ca and Li — with which tricomplex systems are also possible — this complication can left out of account.

b. General condition for the production of tricomplex systems

It is characteristic of tricomplex systems that here the specific charge elements of the colloid ions and micro ions taking part play a part to a very large extent. Before we go into this more fully a general consideration may first be put forward in which we take the above given simple working hypothesis as our starting point.

First of all one must bear in mind then that one can also join the ions which occur in the symbol of the tricomplex systems to the two schemes for the uni- and dicomplex systems.

Compare Fig. 51 in which this is expressed formally as a reaction equation. We can leave aside the question whether the complex relations a of the unicomplex system, as also those b of the dicomplex system are sufficiently intense to make these systems capable of a separate existence — what is meant is a distinct separated phase next to an equilibrium liquid.

In the tricomplex system besides b there occur two other complex relations c and d and it is thereby clear that as regards the question whether in a given case a tricomplex system will be produced or not, it will now depend on the intensity of the complex relations in the tricomplex system compared with those in the unicomplex and dicomplex systems together.

If $2c + 2d + 2b \rangle 2a + 4b$ or $c + d \rangle a + b$, the tricomplex system should be able to form at the expense of the uni and the dicomplex system.
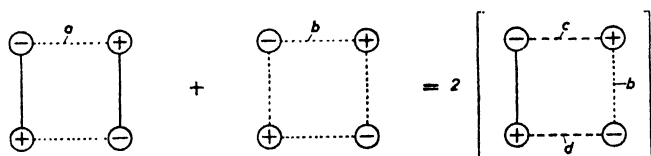
Fig. 51.   Formation of a tricomplex system from (or decomposition to) a unicomplex and a dicomplex system. (see text).

If $c + d \langle a + b$ then conversely the tricomplex system would split up into a unicomplex + a dicomplex system. In this one should bear in mind that the schematic pictures have only formal significance and merely depict the complex relations between four charges while in reality these patterns ought to be applied to the actually existing charge distribution in these systems, for example, to a linear arrangement along the kinked macromolecule distributed in space.

In any case it becomes apparent that the realisation of tricomplex systems will be favoured when the complex relations c and d are more intense than a and especially than b.

If a complex relation indicated by dots means that this is weak, a complex relation indicated by dashes that this is strong, then "reaction equation" of Fig. 51 has at once been adapted to that mode of production of complex flocculations which we have described in § 6 a.

Let us take as example:

isoelectric gelatin + $Mn(NO_3)_2$ + K-chondroitin sulphate.
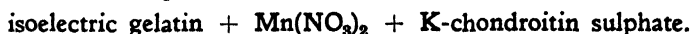
The isoelectric gelatin sol contains amphoions in solution, that is to say, the intermolecular complex relations a are therefore so weak that, given the remaining

forces which bring this macromolecule and the solvent together, they are not able to cause the separation of the gelatin as a unicomplex system.

This holds just as much for the dicomplex combination Mn-chondroitin sulphate. In fact no flocculation or coacervation is produced on mixing $Mn(NO_3)_2$ and K-chondroitin sulphate. It follows from this that b also is weak.

The fact that complex flocculation does occur on mixing isoelectric gelatin, $Mn(NO_3)_2$ and K-chondroitin sulphate must, according to the above given argument, mean that $c + d$ is considerably greater than $a + b$, which in view of the smallness of a and b seems not improbable.

However besides $Mn^{..}$ we also add $NO_3^{.}$ and besides chondroitin sulphate also $K^{.}$, that is to say, in the final mixture there are two different cations and two different anions present besides the amphoion. These can formally furnish as many as four dicomplex systems. (Mn nitrate itself, K chondroitin sulphate itself, KNO_3 and Mn chondroitin sulphate). For all these four b is small which appears from the use of solutions of the first two and further from it remaining clear after mixing. By combination of these four dicomplex systems with gelatin amphoions four tricomplex systems can formally also be produced according to the reaction equation of Fig. 51. We can immediately exclude three of them, because gelatin does not give a precipitate with either $Mn(NO_3)_2$, or K chondroitin sulphate, or KNO_3, so that the tricomplex flocculation in question must contain gelatin amphoions, Mn cations and chondroitin sulphate anions as essential partners. That is to say, only in this latter combination is $c + d$ sufficiently great (see Fig. 52A).

Since however b must be very small, it is actually to be expected that the production of the complex relation between chondroitin sulphate anion and the positive ionised group of the gelatin amphoion is not causally connected with the simultaneous existence of the complex relation between the manganese cation and the negative ionised group of the amphion. In other words it will be present approximately equally strongly after mixing gelatin with K chondroitin sulphate sol.
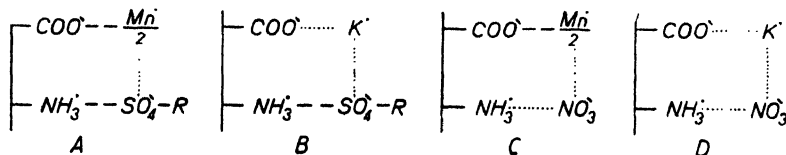
Fig. 52. The four tricomplex combinations which can formally be formed from isoelectric gelatin K chondroitin sulphate (schematically symbolized by $R-SO_4K$; see however p. 420) and $Mn(NO_3)_2$. Intense complex relations are indicated by ————, negligible ditto by ...... A: insoluble tricomplex: B, C and D are soluble since complex relations c and d of Fig. 51 are not simultaneously sufficiently intense in them.

The strongly bound cation is missing in this tricomplex combination (see Fig 52B).

It does not separate out as a independent phase but must be present in the dissolved state in which the K ion is present practically freely dissociated, on the other hand the gelatin amphoion occurs bound together with the chondroitin sulphate anion.

In the clear mixture of isoelectric gelatin sol and $Mn(NO_3)_2$ one must similarly assume a relatively strong complex binding between Mn cation and gelatin amphoion, while the NO_3 ion is practically free (Fig. 52C) and finally in the clear mixture of gelatin with KNO_3 a potential tricomplex is present in which the binding of both cation and anion has practically no longer any complex nature (Fig. 52D).

c. Significance of the colloid equivalent weight and of the composition of the ionised group of the colloid anion

The above formulated condition $c + d \rangle a + b$ is not the only one since $c + d$ itself must also exceed a certain minimum value if one is in fact to arrive at the separation of a tricomplex system. The great significance of a small equivalent weight for the realisation of dicomplex flocculations or coacervations (see p. 374, § 2r and

p. 392, § 3d) might also be expected here. When the condition $c + d \rangle a + b$ is satisfied, the tendency to tricomplex flocculation of the type: amphoion + micro cation + colloid anion would thus increase with decreasing equivalent weight of the participating colloids.

However, we note that carrageen combined either with isoelectric gelatin, or with egg lecithin, has a greater tendency to tricomplex flocculation than chondroitin sulphate (equivalent weights of the colloid anions nearly the same: 271 and 280). This is manifest by the former giving complex flocculation with the divalent cations $Mg^{..}$, $Ca^{..}$, $Sr^{..}$, $Ba^{..}$ and the monovalent $Li^{.}$, while the latter is no longer able to do so.

What is of fundamental significance for tricomplex flocculation is obviously the composition of the ionised group of the colloid anion. For this reason we have compared carrageen (all ionised groups are sulphate groups) with chondroitin sulphate ($50°/_o$ sulphate groups + $50°/_o$ carboxyl groups, not accounted for in Fig. 52).

If one compares together the intensity of the tricomplex flocculation in mixtures: egg lecithin + carrageen + Ca and egg lecithin + pectate + Ca, then this is much smaller in the latter case and the flocculation is extremely small or practically absent in mixtures of egg lecithin + nucleate + Ca.

Now carrageen, pectate and nucleate have not, it is true, the same equivalent weights, but these are at least of the same order of magnitude and the observed differences in intensity of the tricomplex flocculation (strong, weak, absent) can indeed hardly be attributed to this. An explanation is however found in the different polarisability of the ionised groups of these colloids (carrageen $=$ sulphate colloid; pectate $=$ carboxyl colloid; nucleate $=$ phosphate colloid).

This polarisability decreases in the following series from left to right (p. 287, Ch. IX, § 2f):

phosphate group $\rangle$ carboxyl group $\rangle$ water $\rangle$ sulphate group.

Where we wish to bring about a possible tricomplex flocculation with Ca, *i.e.* with a strongly polarising cation (since small and in addition divalent), it will be favourable in all respects that the polarisability of the negative group of the colloid amphoion is large, that of the negative group of the colloid anion small. See p. 418, Fig. 51.

The former leads to a large value of c, the latter to a small value of b, which is favourable for the proposed condition for tricomplex flocculation, namely $c + d \rangle a + b$.

In the combination lecithin + carrageen + Ca this condition is met most completely: the negative ionised group of the amphoion (lecithin) is a phosphate group, that is to say, strongly polarisable.

The Ca ion is small, thus strongly polarising and it therefore concludes powerful complex relations with the phosphate group of the lecithin, that is to say, c is large.

On the other hand the negative ionised group of the colloid anion, carrageen is a sulphate group, that is to say, weakly polarisable and even less polarisable than water. The Ca ion will consequently not want to conclude complex relations with this sulphate group, *i.e.* b is negligibly small.

The complex relation d does not however become small in intensity by the choice of a sulphate colloid as colloid anion. The positive group of the lecithin is comparable with a voluminous organic cation, which (because it cannot have a polarising action, see p. 300, Ch. IX, § 4) will conclude strong complex relations with the carrageen as a colloid of low equivalent weight. Since in addition a is small (the egg lecithin sol is practically stable) the conditions for tricomplex flocculation are therefore very favourable $(c + d \rangle\rangle a + b)$.

If we replace the carrageen by pectate, then these conditions are immediately much less favourable. An unfavourable factor is straight away that b is no longer negligible but on the contrary is already fairly large. The pectate sol with Ca itself already gives a dicomplex flocculation.

Nevertheless the affinity of Ca for the phosphate group of the lecithin is still appreciably greater than for the carboxyl group of the pectate ($c > b$). Formation of a tricomplex system is therefore still possible but this manifests itself only in an intensification of the turbidity, in the presence of lecithin.

In the combination lecithin + nucleate + Ca the factors are now no longer at all favourable for the formation of a tricomplex system.

Here also Ca with nucleate results in the separation of a dicomplex system but this turbidity is hardly, or not at all, intensified in the presence of lecithin. This is also not surprising since b is no longer greater than c; indeed the nucleate and the lecithin have now the same ionised group (phosphate group).

We have as yet discussed tricomplex flocculations in which the amphoion is the association colloid lecithin. It is obvious that the above statements are *mutatis mutandis* also applicable to a protein as amphoion.

Here also the conditions for tricomplex flocculation are favourable in the combination with densely charged sulphate colloids and strongly polarising cations.

On the contrary the combination with densely charged carboxyl colloids and strongly polarising cations is already much less favourable. Indeed the negative ionised group of the amphoion is now also a carboxyl group. The consideration that the polarisability of the — COO′ — groups of the protein is nevertheless stronger than that of the carboxyl colloids (see Chapter IX, § 3, p. 297) offers some prospect of the possibility, though only small, of realising this kind of tricomplex systems. The prospects are quite unfavourable for tricomplex systems from protein amphoions, phosphate colloids (nucleates) and small, strongly polarising cations, because then the polarisability of the negative group of the amphoion is no longer greater than that of the negative group of the colloid anion.

d. Specific ion sequences in the production of tricomplex flocculation, variant colloid amphoion + micro cation + micro anion

If one chooses a particular colloid amphoion (for example egg lecithin or iso-electric gelatin) and investigates with the aid of a number of cations and anions in what combinations of the latter ions complex flocculation occurs or not, one obtains a collection of data from which it is immediately clear that specific properties of the ions play a very great part. See Table 1 and 2 (p. 422 and 423).

In these tables the cation taking part in the tricomplex flocculation is furnished by salt I, similarly the anion by salt II.

The cations of salt I are arranged in valency groups (the divalent in addition in subgroups), while within each division the order from top to bottom is that for increasing reversal of charge concentrations in the "ion spectrum" of egg lecithin (see p. 281, Fig. 10; p. 282, Fig. 12). This order is therefore that of decreasing affinity for the phosphate group.

In reality the order in Table II ought to be that for increasing reversal of charge concentrations in the ion spectrum of the COO′ group of the proteins. This is however not known completely (see p. 297) but in view of the fairly strong

polarisability of this group it may be expected that the order will for the most part fit that of the phosphatides (the order Li — Na — K is the same here for both and transition series occur also with both in the order of the alkaline earth metals).

The cases, in which a tricomplex flocculation was observed with certainty, are indicated in Tables 1 and 2 by a cation symbol in columns 2—6. Here therefore a precipitate is only produced in the mixture of the sol with salt I + salt II, while in each of the three binary combinations (sol + salt I; sol + salt II and salt I + salt II) no precipitate is produced. In addition suppression of the flocculation can be observed in all these cases on addition of KCl (or KNO_3 in the combination with $Pb(NO_3)_2$).

If no flocculation occurs (when the outcome of the three above mentioned control experiments is also negative) in the combination sol + salt I + salt II, then in place of the cation symbol a — is inserted in columns 2—6, that is to say, there is no tricomplex flocculation.

TABLE 1

TRICOMPLEX FLOCCULATION OF EGG LECITHIN SOL WITH A MIXTURE OF TWO NEUTRAL SALTS

Salt I	Salt II.				
	$(NH_4)_6Mo_7O_{24}$	$K_3Fe(CN)_6$	$K_3Co(CN)_6$	K_2HgI_4	KCNS
$Ce(NO_3)_3$	?	Ce	—	Ce	Ce
$La (NO_3)_3$	?	La	La	La	La
$Co(NH_3)_6Cl_3$	?	—(?)	?	$Co(NH_3)_6$	—
$UO_2(NO_3)_2$	?	UO_2	UO_2*	UO_2	UO_2
$Pb(NO_3)_2$	?	Pb	Pb	?	?
$Cd(NO_3)_2$	Cd	?	?	?	Cd
$Mn(NO_3)_2$	—**	?	?	Mn	—
$Cu(NO_3)_2$	Cu	?	?	?	?
$Zn(NO_3)_2$	Zn	?	?	Zn	—
$Co(NO_3)_2$	Co	?	?	Co	—
$Ni(NO_3)_2$	Ni	?	?	Ni	—
$CaCl_2$	Ca	Ca	Ca	Ca	—
$MgCl_2$	Mg	Mg	Mg	Mg	—
$SrCl_2$	Sr	Sr	Sr	Sr	—
$BaCl_2$	?	Ba	Ba	Ba	—
$AgNO_3$	?	?	?	?	—
LiCl	Li	Li	Li	Li	—
NaCl	—	—	—	Na	—
NH_4Cl	—	—	—	NH_4	—
KCl	—	—	—	K	—

* On mixing the two sufficiently dilute salt solutions a light precipitation is produced. Where strong flocculation occurs immediately in the combination sol + salt I + salt II and is reversible, this combination is reckoned as positive.

** This combination occupies a special position since we suspect a reaction between salt I and salt II. The negative outcome of tricomplex flocculation need not therefore cause any surprise.

In the cases in which a simple decision on the occurrence or otherwise of tricomplex flocculation is not possible, since a precipitate has already been produced in the binary combination salt I + salt II, a question mark is inserted in columns 2 — 6.

We may now turn to the discussion of Tables 1 and 2 and for the purpose consider the separate vertical columns. The order in which the salts II have been placed is that from left to right of decreasing anion valency. In what follows we assume

that the effective ainon in K_2HgI_4 is the HgI_4 anion although it remains open to discussion whether it is not the monovalent HgI_3. In either case it appears from the table that the valency of the anion is not exclusively the factor which determines the tendency to produce tricomplex flocculation. We should then have to expect an increasing number from left to right of — signs at the bottom of the columns, while the K_2HgI_4 with di (or mono)valent anion tends much more to tricomplex flocculation than the anions of higher valency situated more to the left in the table.

Also among the monovalent anions clear specific differences of the anions occur. Tricomplex flocculations with egg lecithin or gelatin in which monoatomic anions (I', Br', Cl') take part are absent[1]. The same also holds for NO'_3, on the other hand a certain number occurs with CNS'.

TABLE 2

TRICOMPLEX FLOCCULATION OF ISOELECTRIC GELATIN SOL WITH A MIXTURE OF TWO SALTS

Salt I	Salt II				
	$(NH_4)_6Mo_7O_{24}$	$K_3Fe(CN)_6$	$K_3Co(CN)_6$	K_2HgI_4	K C N S
$Ce(NO_3)_3$	?	Ce	Ce	Ce	Ce ††
$La(NO_3)_3$	?	La	La	La	La ††
$Co(NH_3)_6Cl_3$	?	— (?)	?	$Co(NH_3)_6$	—
$UO_2(NO_3)_2$	?	UO_2 †	UO_2* †	UO_2	UO_2 ††
$Pb(NO_3)_2$	?	Pb	Pb	?	?
$Cd(NO_3)_2$	Cd	?	?	?	Cd
$Mn(NO_3)_2$	—**	?	?	Mn	—
$Cu(NO_3)_2$	Cu	?	?	?	?
$Zn(NO_3)_2$	Zn	?	?	Zn	—
$Co(NO_3)_2$	Co	?	?	Co	—
$Ni(NO_3)_2$	Ni	?	?	Ni	—
$CaCl_2$	Ca	—	—	Ca	—
$MgCl_2$	Mg	—	—	Mg	—
$SrCl_2$	Sr	—	—	Sr	—
$BaCl_2$	?	—	—	Ba	—
$AgNO_3$	?	?	?	?	?
LiCl	—	—	—	Li	—
NaCl	—	—	—	Na	—
NH_4Cl	—	—	—	NH_4	—
KCl	—	—	—	K	—

* and ** see Table I, † Reversible with NaCl, when very dilute UO_2 nitrate is used. †† Large excess of salt I is necessary here.

This already reflects the affinity sequence of these monovalent anions with respect to colloid cations. We may draw attention to the anion spectra (order of the reversal of charge concentrations) of positive proteins (Fig. 22, p 299), in which of the anions mentioned the CNS' always possesses the lowest reversal of charge concentration.

---

[1] Or escape observation by the method employed here (sol + salt I + salt II) if the bonds in the tricomplex system are very weak because the two ions of salt I and salt II, which do not take part in tricomplex formation (NO_3' or Cl' of salt I and K· of salt II) form an indifferent salt which suppresses a very weak complex flocculation. Compare small print on p. 432 from which it appears that such a very weak tricomplex flocculation is possible with phosphatide ⊢ Ca·· ·⊦ I'.

Nevertheless this affinity of the CNS' is still relatively weak, as appears from the fact that tricomplex flocculation is only realisable with egg lecithin and gelatin by means of micro cations which have a very strong polarising action (for example $Ce^{...}$, $La^{...}$, $UO_2^{..}$, $Cd^{..}$) and therefore are situated far to the left in the reversal of charge spectra of the phosphatides (Fig. 12 and 20, p. 282 and p. 295). This is in agreement with the condition $c + d \rangle a + b$ discussed in § 6b (p. 418). A relatively small intensity of the complex relation d (binding of CNS' to the amphoion) can be compensated by a very intense complex relation c (binding of the cations mentioned to the amphoion).

In combinations with Mo_7O_{24}'''''', $Fe(CN)_6'''$, $Co(CN)_6'''$ and HgI_4'' tricomplex flocculation occurs even with the micro cations which have a smaller affinity for the negative group of the amphoion than the above mentioned one. Thus this still takes place in the case of egg lecithin with Mg, Ca and Sr[1]. Isoelectric gelatin also gives tricomplex flocculations with molybdate or with HgI_4'' in the case of Mg, Ca, Sr and Ba but not with $Fe(CN)_6'''$ or $Co(CN)_6'''$.

These data are also in agreement with the above. One can in fact reasonably expect that the four anions mentioned will have a greater affinity for the positive group of the colloid anion than the CNS' ion. Indeed these voluminous anions are pronounced alkaloid reagents and strong protein precipitants in acid media, that is to say very readily produce insolubility in combination with voluminous cations. For these reasons the affinity for the positive group of the amphoion must therefore be likewise great (d large). The condition $c + d \rangle a + b$ anticipates then that even a smaller c suffices to lead to tricomplex flocculation. A smaller c means that now cations which are situated more to the right in the reversal of charge spectra of egg lecithin or negative gelatin also give tricomplex flocculation, such as the above mentioned alkaline earth metals.

Of the four anions mentioned the HgI_4'' must have the largest d value by far, since it also gives tricomplex flocculation with $Li^.$, $Na^.$, $NH_4^.$ and $K^.$, both with gelatin and with egg lecithin.

It is further interesting that the remaining three anions still give tricomplex flocculation in the case of egg lecithin with $Li^.$ but not with the other alkali cations.

This point of detail can also be anticipated: for these anions d is smaller than for HgI_4''. If therefore tricomplex flocculation is still possible with an alkali cation, the most strongly polarising one, i.e. $Li^.$, is the most promising one.

The three anions mentioned no longer give tricomplex flocculation in the case of isoelectric gelatin with Li, which is not surprising since the polarisability of the carboxyl group of the gelatin amphoion is smaller than that of the phosphate group (of egg lecithin).

Finally we discuss the tricomplex flocculation with $Co(NH_3)_6^{...}$. We cannot expect any polarising action for this voluminous complex ion, in contradistinction to the small (likewise trivalent) $Ce^{...}$ and $La^{...}$. The intensity of the complex relation c is consequently much smaller with this complex trivalent cation than with $Ce^{...}$ or $La^{...}$[2]. It need thus not surprise us that in the combinations with $Fe(CN)_6^{...}$

---

[1] Since ammonium molybdate already gives a precipitate with Ba itself in the control experiment, it is not possible to establish whether tricomplex flocculation occurs in the combination colloid amphoion + Ba + paramolybdate.

[2] The much smaller affinity of polyvalent complex cations for the negative ionised group is also already manifest in the study of ion antagonism. Compare the situation apart of the hexol$^{......}$ and the $Co(NH_3)_6^{...}$ in Fig. 34, p. 313.

and CNS' with $Co(NH_3)_6$··· no tricomplex coacervation occurs, although this does occur with Ce··· and La··· and even with divalent small cations (and in the case of egg lecithin even with the monovalent Li ion).

A clear tricomplex flocculation occurs however with $Co(NH_3)_6$··· in the combinations with HgI_4'' but we already saw that this anion tends most to tricomplex flocculation so that here even the monovalent cations, among them even the most unfavourable, K·, give tricomplex flocculation.

e. Specific cation sequence in the production of tricomplex flocculation, variant: colloid amphoion + micro cation + colloid anion

In the preceeding subsection we have spoken without further comment of the greater or less intensity of complex relations between micro ions and colloid ions. Naturally what was meant was actually the comparison of the maximum intensities of these complex relations obtainable with the micro ions in question. Indeed, just as this is already the case in the dicomplex systems, colloid cation + micro anion or colloid anion + micro cation, the statement, that the intensity of the complex relations formed by a micro cation is still a function of the concentration, also holds for the tricomplex systems.
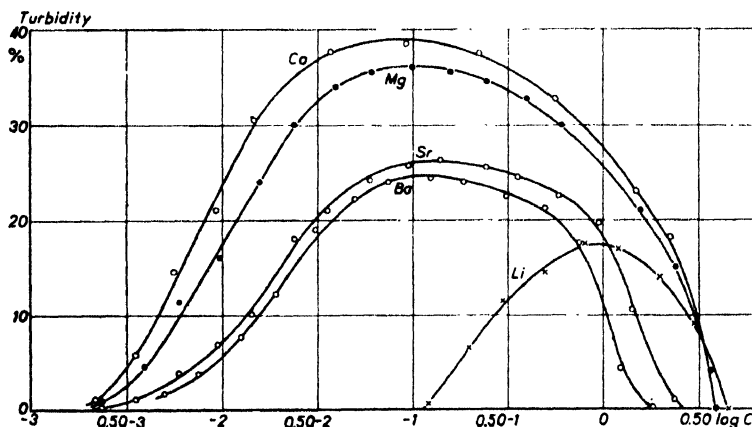
Fig. 53. Comparison of some cations as regards the tricomplex flocculation of a mixture of of soya bean phosphatide sol and carrageen sol.
The mixing proportion of phosphatide and carrageen was kept constant; the cations were added as chlorides.
Ordinates: turbidity.
Abscissae: logarithms of the salt concentrations (in eq. per l).

At a suitably chosen ratio of colloid amphoion and colloid anion, the added cation brings about tricomplex flocculation from a certain small concentration onwards. On increasing the cation concentration the flocculation increases to a maximum and then decreases again.

One should compare the general shape of the curves in Fig. 53 which reproduce the turbidity of the tricomplex flocculation after a certain interval of time of alcohol

soluble soya bean phosphatide + carrageen + certain cations (added as chlorides) [1].

The same consideration, as was discussed in the dicomplex flucculation, variant colloid anion + micro cation (p. 388 and 391), also holds here:

Along the left-hand rising branches only, the cation of the salt is predominantly active. It initiates and strengthens the new complex relation between negative ionised group (here of the amphoion) and the cation. The maximum of the complex relation is reached when finally the salt concentration becomes so great that the general suppressive action gradually becomes noticeable. The branch falling towards the right is therefore an auto-suppression of the tricomplex flocculation by just the salt which in the smaller concentrations has lead to the production of the tricomplex flocculation. In this both the cation and the anion of the salt used will play a part.

These ideas are confirmed by a comparative investigation of the tricomplex flocculations: egg lecithin + carrageen + various Ca salts. It appears, in fact, that the left-hand rising branches of the curves, to the maximum of the $CaCl_2$ curve, cover each other completely, when one make comparisons between $CaCl_2$, $CaBr_2$, CaI_2, $Ca(NO_3)_2$ and $Ca(CNS)_2$. It therefore does not matter here what anion accompanies the Ca ion. On the other hand the branches of the curves no longer cover each other beyond the maximum of the $CaCl_2$ curve, that is to say, the anion also plays a part in the auto-suppression. We shall return to this in § 6f (see small print on p. 432).
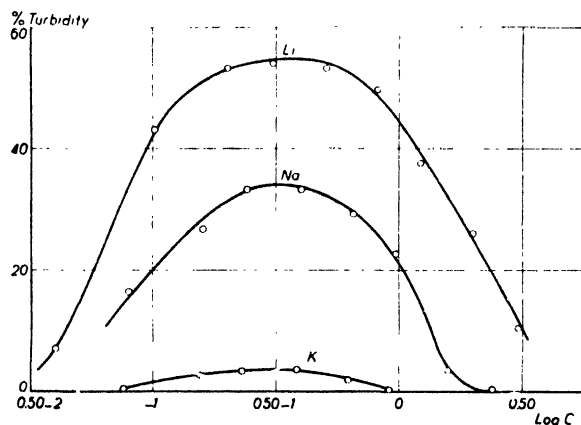
Fig. 54. Comparison of Li, Na and K as regards the tricomplex flocculation of a mixture of soya bean phosphatide sol and carrageen sol.
Mixing proportion of the colloids, abscissae and ordinates as in Fig. 46. To strengthen the tricomplex flocculation the medium contains a constant 25% of alcohol.

We now turn once again to Fig. 53. In this graph we see that all five curves exhibit a maximum. This maximum lies for Li at a higher concentration ($\pm$ 1N) than for the alkaline earth metals ($\pm$ 0.1 N), which is a consequence of the fact that the latter already reach their maximum complex relations at lower concentrations. The mutually very different heights of the five maxima is a matter which specially interests us.

The tricomplex flocculation is the strongest with $Ca^{\cdot\cdot}$ and decreases from $Mg^{\cdot\cdot}$, through $Sr^{\cdot\cdot}$ and $Ba^{\cdot\cdot}$ to $Li^{\cdot\cdot}$.

Now NaCl and KCl give no tricomplex flocculation at all but do so when one

---

[1] H. G. BUNGENBERG DE JONG and C. H. RERING, Proc. Koninkl. Nederland. Akad. Wetenschap., Amsterdam, 45 (1942) 679.
The remaining data in this and the next subsection have been taken from two further publicatoins of the same authors: Proc. Koninkl. Nederland. Akad. Wetenschap., Amsterdam, 45 (1942) 705 and 713.

works not with an aqueous but with a 25% alcoholic medium. Fig. 54 gives a comparison
of the three cations in this medium.

By strengthening of the complex relations in this medium containing alcohol
(see for this action of alcohol or acetone p. 399, 401, 404 and Fig. 47 on p. 403) the
maximum of the LiCl curve is now displaced towards smaller concentrations and in
addition the tricomplex flocculation has become stronger. From this figure it follows
that the intensity of tricomplex flocculation decreases in the order Li $>$ Na $>$ K.

A few other cations have been compared in aqueous medium as regards the height
of the turbidity maximum and these results combined with those above then furnishes
the following series for maximum strength of the tricomplex flocculation:

$$Cd^{..} \ > \ Ca^{..} \ > \ Mg^{..} \ > \ Ag^{.} \ > \ Sr^{..} \ > \ Ba^{..} \ > \ Li^{.} \ (> \ Na^{.} \ > \ K^{.})$$

in which those in brackets no longer give any tricomplex flocculation at all in aqueous
medium but do so in 25%
alcohol.

For a discussion of this
sequence we again start from
the condition $c + d \ > \ a + b$
and since we are working
throughout with the same
amphoion and the same
colloid anion (and these
both in the same ratio in
addition) a and d are con-
stant. Replacement of the
one cation by another only
causes each time different
values of c and b. It is
therefore the ratios c/b
which will determine the
position of a cation in the
given sequence.

Electrophoretic mea-
surements on the phos-
phatide alone and on the
carrageen alone now give
support to this. See Fig. 55
and Fig. 56 in which the
measured electrophoretic
velocities are plotted as a
function of the logarithm of
the salt concentration. We
see that for each ion the
relation holds that the re-
versal of charge concentr-

Fig. 55. Reversal of charge curves for phosphatide and
carrageen sols with LiCl, NaCl and KCl. U: electrophoretic
velocity in arbitrary units; log C: logarithms of the salt con-
centrations (eq. per l). For bend in the left KCl curve, see
note on p. 333.

ation is lower with the phosphatide than with the carrageen. A lower reversal of
charge concentration means a greater affinity.

Thus for all ions the condition $c \ > \ b$ is satisfied.

We see in Fig. 55 further that the distance between the two reversal of charge points is small for K·, larger for Na· and appreciably still much larger for Li·. This means, since the concentrations are plotted logarithmically in these figures, that the ratio of the reversal of charge concentrations, that is to say, the ratio of c/b increases in the order K· — Na· — Li·, which is in agreement with the differences in intensity
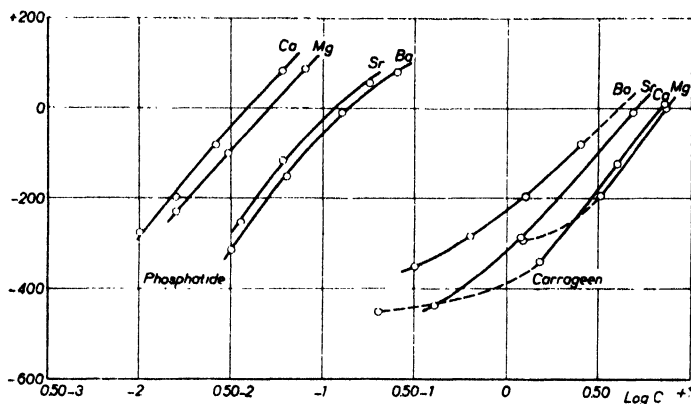
Fig. 56. Reversal of charge curves for phosphatide and carrageen sol with $MgCl_2$, $CaCl_2$, $SrCl_2$ and $BaCl_2$.
Ordinates and abscissae as in Fig. 55. For bends in the right $CaCl_2$ and $MgCl_2$ curves see note on p. 333.

of the tricomplex flocculation in 25% alcohol (see p. 426, Fig. 54). If one reads off these logarithmic distances of the reversal of charge points for the other ions also (and also for the ions Cd·· and Ag· not included in the figures 55 and 56) it is seen that these distances arrange themselves in the series:

$$Cd·· \rangle Ca·· \rangle Mg·· \rangle Ag· \rangle Sr·· \rangle Ba·· \rangle Li· \rangle Na· \rangle K·$$

i.e. in the same series order as given above (p. 427) for the maximum strenghts of the tricomplex flocculation.

See also Fig. 57 in which the maximum turbidities in aqueous medium are plotted against the logarithmic differences in reversal of charge concentrations of the cations indicated.

We thus see that with K· and Na· the ratio c/b is still too unfavourable in aqueous medium to cause tricomplex flocculation and that with the remaining cations the intensity of the flocculation rises with the value of c/b.

Similar results in principle have been obtained in the combinations egg lecithin + carrageen + micro cation and isoelectric gelatin + carrageen + micro cation.

In the last mentioned combination tricomplex flocculation occurs in aqueous medium not only with the alkaline earth cations and Li, but also with Na· although weakly. The K· ion is here also again the most unfavourable cation.

It is further interesting that this tricomplex flocculation is not restricted only to the isoelectric point of the gelatin used (pH 5.1), but still also occurs at higher pH values (pH 6 — 8) and disappears only at about pH = 10.

This is possibly of importance for biology since it appears from it that at pH values of 6 — 8, the so-called acid proteins can still react sufficiently as amphoions (multipole ions) to make it possible that tricomplex systems could occur in the conditions of natural media.

Fig. 57. Correlation between intensity of the tricomplex floc-culation (each time under optimum conditions) with nine cations and the logarithmic dif-ference of the reversal of charge concentrations of the colloids with these cations.
Ordinates: height of the maxi-mum of the turbidity curves (for example in Fig. 53).
Abscissae: the logarithm of the quotient of the reversal of charge concentrations (the logarithmic difference in, for example, Fig. 55 and 56).
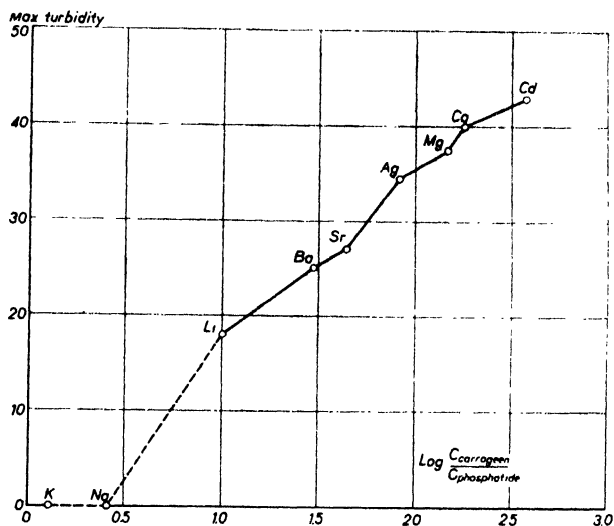
f. Specific ion sequences in the suppression of tricomplex flocculations

If one has brought about a tricomplex flocculation, for example, phosphatide + carrageen + Ca und now one adds extra salts which cannot themselves bring about tricomplex flocculation, then the tricomplex flocculation is suppressed.
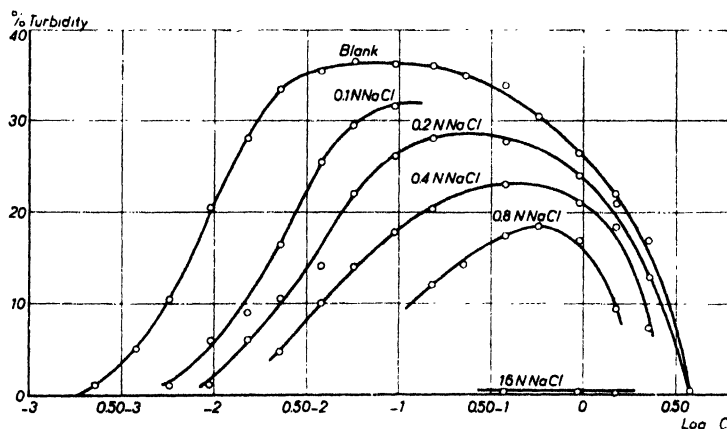
Fig. 58. Influence of a number of constant NaCl concentrations on the tricomplex flocculation of soya bean phosphatide + carrageen + $CaCl_2$.
Ordinates: turbidity.
Abscissae: logarithms of the $CaCl_2$ concentrations (eq. per l).

Compare Fig. 58 in which blank represents the flocculation by $CaCl_2$ only, and the remaining curves the effect of $CaCl_2$ in the presence of successively higher constant NaCl concentrations. NaCl of 1.6 N thus suppresses the flocculation completely at all $CaCl_2$ concentrations.

In such suppressions specific sequences of both cations and anions make their appearance in a very pronounced manner.

For the investigation we choose a constant $CaCl_2$ concentration such that we are at the maximum of the blank curve. The results for the tricomplex flocculation egg lecithin + carrageen + Ca are reproduced in Fig. 59A and Fig. 60A.
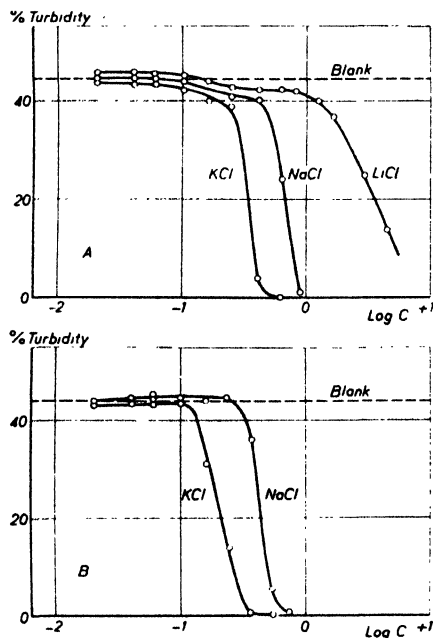
Fig. 59. Specific differences of the cations in the suppression of the tricomplex flocculation of egg lecithin + carrageen + Ca or Li with chlorides of the type 1—1.
A: in the tricomplex with Ca.
B: in the tricomplex with Li.
Ordinates: turbidity.
Abscissae: logarithms of the salt concentrations (eq. per 1).
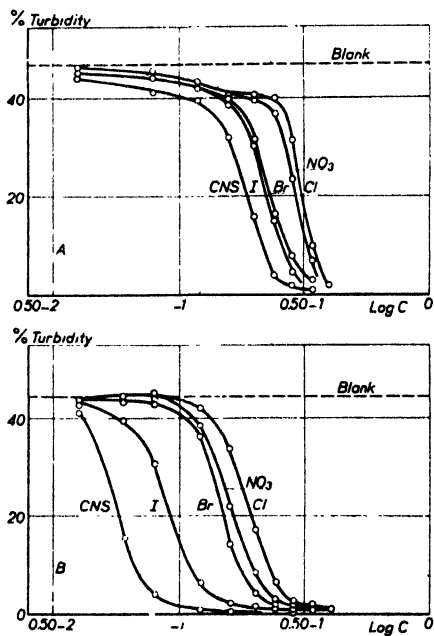
Fig. 60. Specific differences of the anions in the suppression of the tricomplex flocculation of egg lecithin + carrageen + Ca or Li with K salts of the type 1 — 1.
A: in the tricomplex with Ca.
B: in the tricomplex with Li.
Ordinates: turbidity.
Abscissae: logarithms of the salt concentrations (eq. per 1).

The suppressive action of monovalent cations thus decreases from left to right in the order: K $\rangle$ Na $\rangle$ Li and that of monovalent anions in the order CNS $\rangle$ I $\rangle$ Br $\rangle$ NO₃ $\rangle$ Cl.

It is striking that the LiCl curve in Fig. 59A lies relatively so very far to the right. But this already becomes understandable by considering that LiCl itself brings about

a tricomplex flocculation in the mixture egg lecithin + carrageen, although a weak one compared with Ca, the top of which lies around log C = O (similar as in Fig. 53, p. 425). In the suppression of the Ca tricomplex flocculation we therefore first pass over into the weaker Li tricomplex flocculation and then actually follow the branch of the curve of the Li tricomplex flocculation descending towards the right (compare the analogous case on p. 398, Fig. 43 for x = 8)

The same cation and anion sequences also occur in the suppression of Li tricomplex flocculation. See Fig. 59B and 60B.

In the interpretation of these specific sequences we shall again have to take as starting point the affinity sequences such as appear in the "ion spectra" (p. 285, 295). Further it should be borne in mind that matters are more complicated here than in the similar discussion regarding the specific sequence in the suppression of dicomplex coacervation or flocculation. There (p. 376) we had only one complex relation to deal with, here with two present simultaneously[1]:

I. phosphate group of the egg lecithin ..... Ca.
II. quaternary ammonium group of the egg lecithin .... carrageen anion.

In addition one needs to consider whether the suppressive ion at the concentrations at which it is used, brings about practically no new complex relations still or very weak ones, or whether these already attain a considerable value. In the first case the order of these ions will be that in which they occur in the "ion spectrum", in the latter case the order will be just the reverse (compare p. 391 and 398).

This latter consideration quite clearly plays a part in the interpretation of the cation sequence found, K ⟩ Na ⟩ Li (Fig. 59).

These cations acting on II create no new complex relations, they do not flocculate carrageen. Therefore one must expect in this case the sequence K ⟩ Na ⟩ Li, i.e. the sequence in the ion spectrum of carrageen (p. 285, Fig. 15).

The alkali cations however also act on I, but they, at least Li and Na, bring about new complex relations. In fact Li itself can already bring about a tricomplex flocculation and Na is very close to this (gives tricomplex flocculation on addition of 25% alcohol).

Therefore one must expect in this case the opposite order from that in the ion spectra of phosphate colloids (p. 295, Fig. 20). This latter is Li ⟩ Na ⟩ K. We should thus expect K ⟩ Na ⟩ Li. The result of these arguments is that the alkali cations acting both on I and on II in the suppression of tricomplex flocculation will arrange themselves in the order K ⟩ Na ⟩ Li, which is in agreement with experiment.

As regards the anion sequence, we shall be able to neglect the influence on I (being that on Ca), so that there remains the action on II, being that on the quaternary ammonium group of the egg lecithin. In view of the fact that this group can be regarded as a large and thus non-polarising, cation, we must expect here the same affinity sequence as in the ion spectra of positive proteins (Fig. 22, p. 299), that is I ⟩ Br ⟩ Cl for the monatomic ions and for the others CNS ⟩ NO_3, in which CNS as a rule comes before I, NO_3 usually as the left-hand or right-hand neighbour of Br.

The concentrations at which CNS had a suppressive action, bearing in mind the small Ca concentrations, are still so small that there is no question of a separation

---

[1] The third, Ca — carrageen, corresponding to *b* in the scheme, Fig. 51 (p. 418), can be left out of consideration since it is very weak or absent.

of a new tricomplex system egg lecithin $+$ Ca $+$ CNS[1]. We therefore expect for the anion sequence that of the affinity sequence of the positive group of the egg lecithin. The sequence experimentally found (p. 430, Fig. 60):

CNS $\rangle$ I $\rangle$ Br $\rangle$ NO$_3$ $\rangle$ Cl, is in agreement with this.

In the suppression of the tricomplex flocculation, gelatin $+$ carrageen $+$ Ca, the same specific ion sequence of the monovalent cations and anions was also found. For the interpretation we must go through the same arguments again bearing in mind that the two complex relations are now:

I. carboxyl group of gelatin . . . . . Ca

II. basic groups of gelatin . . . . carrageen anion.

Since the sequence is also Li $\rangle$ Na $\rangle$ K in the ion spectrum of negative proteins and these ions bring about fairly strong complex relations in this case also, the interpretation with regard to the cation sequence in the suppression of the tricomplex flocculation is exactly the same as that described in detail above. For the anion sequence, arguing similarly as above, the ion sequence in the ion spectrum of gelatin (p. 299, Fig. 22) is to be expected. This is indeed the same as that found experimentally in the suppression of the tricomplex.

The general shape of the curves such as occur in Fig. 53 (p. 425) was discussed in § 6e. In the example used there (egg lecithin $+$ carrageen $+$ Ca) it was found that the anion accompanying the Ca ion does not matter for the left-hand rising branch of the curve but that the accompanying ion does matter in the branch descending to the right. The latter branch as stated there actually represents an auto-suppression of the tricomplex flocculation and the specific differences which occur here on a comparison of CaCl$_2$, CaBr$_2$, CaI$_2$, Ca(NO$_3$)$_2$ and Ca(CNS)$_2$ are only interpretable after the matters discussed in this subsection. The run of the curves is fairly complicated and we do not propose to go into details here. It may suffice to state that especially the curve for Ca(CNS) is very anomalous and to a lesser extent that for Ca I$_2$. It is of interest now that egg lecithin gives a fairly weak tricomplex flocculation of the type amphoion $+$ micro cation $+$ micro anion with Ca (CNS)$_2$ (in the absence of carrageen) and also with Ca I$_2$ although very weak[2]. The remaining Ca salts do not do this.

It need therefore cause no surprise that the sequence of these descending branches of the curves is not that discussed above, but that the CNS lies quite anomalously at the end:

I $\rangle$ Br $\rangle$ NO$_3$ $\rangle$ Cl $\rangle$ CNS.

Furthermore it was observed that on increase of the concentration the curve for CaI$_2$, which at first suppresses the most strongly, is also displaced to the right:

Br $\rangle$ I $\rangle$ NO$_3$ $\rangle$ Cl $\rangle$ CNS.

These displacements of CNS and of I from left to right are analogous to those discussed in § 3c (p. 391).

---

[1] See small print on this page.

[2] These tricomplex flocculations are really the simplest of those discussed so far because beside the amphoion only those ions are added which occur essentially in the tricomplex. With all the others we had, besides the amphoion, four ions, namely two cations, two anions. That is two too many, which thus can only have in principle a suppressive action. It is probable for these reasons that in Table I (p. 422) where we have investigated egg lecithin $+$ salt I $+$ salt II there was no evidence for the existence of the tricomplexes egg lecithin $+$ Ca $+$ CNS or egg lecithin $+$ Ca $+$ I.

XI. MORPHOLOGY OF COACERVATES

H. G. Bungenberg de Jong

Department of Medical Chemistry, University of Leyden

§ 1. STATIC COLLOID MORPHOLOGY

a. Introduction

The tasks of colloid morphology and its two subdivisions, static and dynamic colloid morphology have already been expounded in brief outline in chapter I.

In this chapter we shall restrict ourselves to the simplest microscopic objects of colloid morphology, namely to the cases in which the colloid body consists of one or more coacervates. In addition colloid bodies will be considered which although they no longer consist of coacervates have been produced from them.

In this first § we deal mainly with morphological systems in equilibrium (static colloid morphology), in the succeeding §§ morphological changes of state (dynamic colloid morphology) are more prominent.

For the explanation of the properties (shape, changes of state, etc.) of a colloid body of microscopic[1] dimensions which consists of a limited amount of coacervate surrounded by equilibrium liquid, one has in general to take into account 1. the properties of the three-dimensional contents and 2. those of the two-dimensional surface. The first mentioned properties (e.g. viscosity, composition) can be studied on coacervates in bulk and these have already been described in previous chapters (VIII and X). The properties of the two-dimensional surface are equally important. The electrophoretic charge of complex coacervates drops has already been discussed in chapter X § 2e (p. 345) but for this chapter it is the interfacial tension which is especially of interest.

That adequately liquid coacervates distributed in microscopic amounts assume the spherical shape is thus the consequence of the properties of the three-dimensional contents — the liquid nature of the coacervate — and of the two-dimensional surface — the presence of an interfacial tension coacervate/equilibrium liquid.

b. Interfacial tension coacervate/equilibrium liquid

In many of the succeeding sections wetting plays an important part and for the sake of explanation one would be glad to possess further data on the interfacial tensions and contact angles. However we must disappoint the reader on this point so that much of what is offered in this chapter is still only a description of phenomena, while the explanation in general still remains in the back-ground.

The reason for this is that we are still inadequately informed with regard to these

---

[1] This choice reduces considerably the distorting influence of gravity. By the difference in sp. gr. of coacervate and equilibrium liquid a coacervate body of macroscopic dimensions is forced into the shape of the containing vessel (see p. 233, Fig. 3). Such artificial shapes do not interest us in this chapter.

data so desirable for an explanation. That this region is still so little worked rests mainly on technical difficulties as a result of which it was until recently impossible to measure the interfacial tension coacervate/equilibrium liquid. It is true we were already convinced that this interfacial tension must be small in thoroughly liquid coacervates. Indeed when the latter settle in narrow tubes to form a homogeneous coacervate layer the boundary surface coacervate/equilibrium liquid still stands more or less horizontal.

Also attempts to measure this interfacial tension in the complex coacervate gelatin (positive) + gum arabic (negative) with the Du Noüy tensiometer remained without result: the platinum ring was pulled out of the boundary surface even with the slightest force. Success has only recently been achieved with the aid of a microscopic capillary tube method of measuring this interfacial tension, whereby extraordinarily small values were indeed obtained[1].

In this a beginning was made at the same time with the study of the factors which influence this interfacial tension. For so far as we can estimate them already, these are the same factors which also govern the intensity of the complex relations. And indeed it is a fact that a variable which intensifies these complex relations causes the interfacial tension to increase, conversely the factors which cause the complex relations to decrease in intensity similarly makes the interfacial tension decrease.
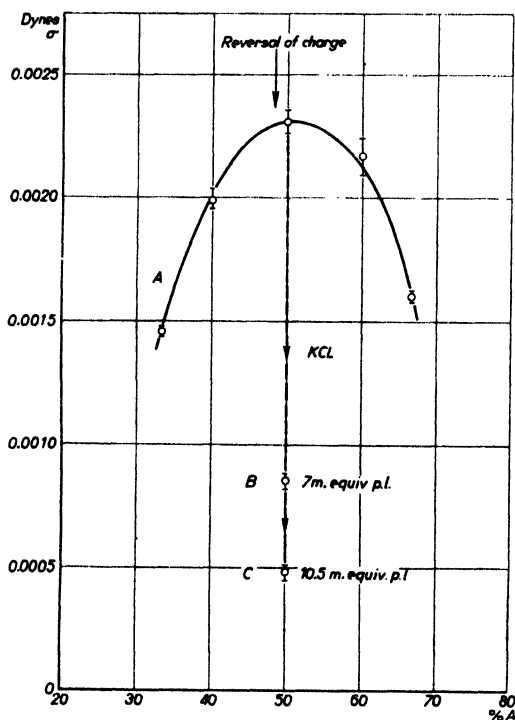
Fig. 1. Influence of certain variables on the interfacial tension coacervate/equilibrium liquid in the complex coacervate gelatin — gum arabic.
Ordinates: Interfacial tension in dynes per cm.
Abscissae: Mixing ratio of the two 2% isohydric sols expressed in % of the gum arabic sol (10 m. eq. per 1 Na acetate and 100 millimol. per 1 acetic acid present as buffer).
Curve A: influence of the mixing ratio on σ; points B and C: interfacial tensions for the mixing ratio 50% A in the presence of 7 and 10·5 m.eq. per 1 KCl. Arrow = position of the reversal of charge point.

Compare the curve A in Fig. 1, which represents the result of a variation of the mixing ratio of isohydric (pH 3·7) 2% gelatin and gum arabic sols. The reversal of charge point was determined electrophoretically as lying at 48% A (indicated by

---
[1] L. DE RUITER and H. G. BUNGENBERG DE JONG, *Proc. Koninkl. Nederland. Akad. Wetenschap.*, Amsterdam, 50 (1947) 836.

the arrow). Now we know however from earlier investigations that the electrophoretic reversal of charge point lies at or close to the mixing ratio at which the complex relations are a maximum (see p. 345 and 357, Ch. X § 2 e and 2 h). We now see from fig. 1 that the interfacial tension exhibits a maximum at the reversal of charge point, which might in fact be expected *a priori* from the theory of LIPPMANN for the electrocapillary curve (see Volume 1).

In Fig. 1 the points B and C represent the interfacial tension for the mixing ratio 50% A in the extra presence of $7 \cdot 5$ and $11 \cdot 5$ m.eq. KCl per litre.

We already know that added indifferent salts reduce the complex relations in intensity (and finally suppress the coacervation, see Ch. X, § 2f and 2l, p. 349 and 364). We are thus not surprised to observe a very considerable lowering effect of the added KCl on the interfacial tension.

Finally we must take into account this lowering influence of salts in a discussion of the already very small interfacial tension at the maximum of the σ curve A . The coacervated system is in fact in no sense free from salts. In the first place buffered A and G sols were used which contained 10 m. eq. p. 1 Na acetate + 100 millimol p. 1 acetic acid.

Further one has to take into account the presence of about 8 m. eq. p. 1 Ca acetate produced from the counter ions of both colloids (see p. 368). The observed maximum value of σ of $2 \cdot 3 \times 10^{-3}$ dynes, is thus undoubtedly appreciably lower on account of the presence of 10 m. eq. p. 1 Na acetate + 8 m. eq. p. 1 Ca acetate than must be expected for a coacervated system in which these salts are not present.

Naturally it is not permissible to extrapolate to a salt content zero, but one would then still arrive at an estimate of the value of the interfacial tension which is always very remarkably low (less than 10^{-2} dynes per cm.).

For a provisional theory of these very low interfacial tensions in complex coacervates see the original publication.

c. Wetting phenomena

Coacervates of all kinds of types, provided the colloid belongs to the hydrophilic type, wet a glass surface and spread over it when they settle on it from the equilibrium liquid. We have already mentioned (p. 339, note 3) that covering the glass surface with a dried layer of amylum solubile prevents this spread as a result of which such starched microscope slides form an indispensible aid in the microscopic study of the colloid morphology of these coacervates.

With respect to this wetting behaviour amylum stands quite by itself. If one smears the slide with vaseline the spread of the coacervate drops is not prevented.

This is connected with the fact that the above mentioned group'of coacervates very generally take up drops of organic liquids which are not miscible with water when these are offerred in the equilibrium liquid[1] (p. 436, Fig. 2a).

They therefore share with glass the property of being completely wetted by the coacervate.

Carbon and carmine particles are also taken up (Fig. 2 b) but amylum grains merely attach themselves to the outside of the coacervate drops. This is in agreement with the difficult wettability of a starched microscope slide by coacervate drops.

---

[1] H. G. BUNGENBERG DE JONG and A. J. W. KAAS, *Bioch. Z.*, 232 (1931) 338.

Other objects of biological origin can also be taken up (pollen grains, various mould spores) or merely be attached to the outside[1].

Coacervates of phosphatides and of oleates behave differently with regard to

a   b
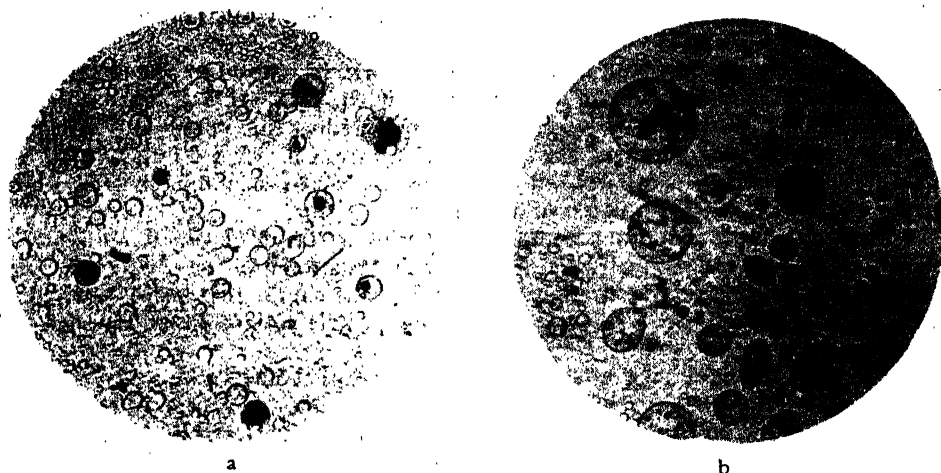
Fig. 2.   Complex coacervate drops of gelatin — gum arabic with included drops of an organic liquid or with included solid particles ($+75$ · lin.).
a: with included tetrahydro naphtalene drops, which are coloured with Sudan III to mark them.
b: with included carbon particles (,,Norit'').

wetting. For example they do not spread on a glass surface and thus can be observed without starched slides. The discrepant behaviour of these coacervates must indeed be connected with the pronounced "lipophilic" properties of these association colloids.

d. Horizontal observation

Drops of the gelatin (positive) + gum arabic (negative) complex coacervate, lying on a starched slide, give the impression of being spherical on account of their perfectly round contours. Viewed with a horizontally placed microscope it appears however that small free coacervate drops (about 100 μ) are already considerably flattened and larger ones approximate more to the shape of discs. Although the sp. gr. of coacervate and equilibrium liquid only differs by a few per cents (about 4%), this small difference is sufficient with the extremely small interfacial tension coacervate/equilibrium liquid (p. 434) to bring about a considerable flattening.

In general it is to be recommended in the study of colloid morphological problems to observe not only vertically but also horizontally and furthermore if possible to study the freely suspended colloid bodies (with horizontal direction of observation).

An inclusion may for example, with vertical observation give the impression of lying centrally in a coacervate drop (Fig. 3A) while horizontal observations proves

---

[1] See H. G. BUNGENBERG DE JONG, *Protoplasma*, 15 (1932) 110, see p. 161—162.

nevertheless that the inclusion reclines on the inside of the coacervate boundary surface:

Fig. 3. Different images of the position of inclusions in coacervate drops.
In this case a spherical object, for example a drop of organic liquid, is considered as the inclusion. With horizontal observation the inclusion, according to its specific gravity lies pressed against the bottom (B) or the top (C) of the coacervate drop. With vertical observation (A) the inclusion appears to take up a central position in both cases.
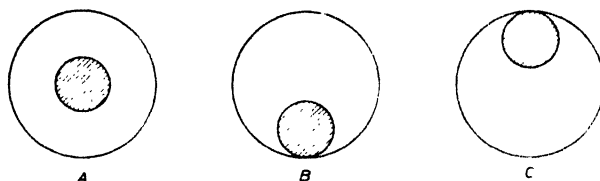
and indeed at the lowest point (inclusion specifically heavier, Fig. 3B) or at the highest point (inclusion specifically lighter, Fig. 3C).

e. Gelatin gel bodies with enclosed gum arabic vacuole

The facts mentioned in § 1 d brings us all the time up against the question whether the inclusion is really completely enclosed by the surrounding coacervate or indeed if a three-phase contact coacervate—inclusion—equilibrium liquid is present. This question cannot be solved by morphological observation alone.

In the first case the inclusion, since it is specifically heavier or lighter than the coacervate, lies pressed against the unobservable thin coacervate lamella.

In the second case it is attached to the coacervate surface with a very acute angle of contact and the whole rolls in such a way that positions B or C are taken up.

Thus the figures 3b and 3c are observed when viewed from the side in both cases.

However success in preparing the colloidal bodies mentioned in the title of this subsection speaks in favour of an angle of contact zero, that is to say of a complete submersion of an organic liquid in the coacervate [1].
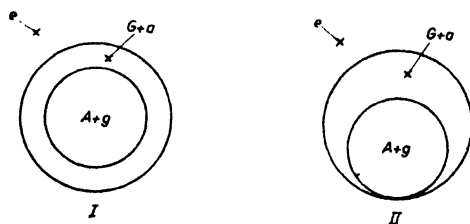
Fig. 4. Scheme to illustrate the production of gelatin gel bodies with enclosed gum arabic vacuole. When warm the gelatin-rich (G + a) liquid completely wets the concave surface of the surrounding emulsifying liquid e, so that the gum arabic-rich (A + g) liquid no longer has any contact with the emulsifying liquid. On cooling the G + a shell gelates, and then surrounds a cavity containing gum arabic in solution.

In this preparation one emulsifies a coacervated system of a mixture of concentrated gelatin and gum arabic sols containing $MgSO_4$ (p. 255, Ch. VIII, § 8), in an excess of a mixture of chloroform and toluene [2]. If the gelatin-rich liquid layer (G + a) is the coacervate and the gum arabic-rich layer (A + g) the equilibrium liquid

---

[1] H. G. BUNGENBERG DE JONG and O. BANK, *Protoplasma*, 33, (1939) 512, see p. 525—527.
[2] To promote the emulsification 1% egg lecithin is also dissolved in it.

then the cavities in the emulsifying liquid must be completely coated by the
gelatin-rich liquid layer (Fig. 4 I) so that the latter completely separates the A + g
liquid from the emulsifying liquid (e). Here also on account of the difference in
sp. gr. the surrounding lamella can be very thin (Fig. 4 II).

If now one cools the preparation the gelatin coacervate membrane gelates and
one can now remove the emulsifying liquid with ether and than transfer the objects
into water.

It then appears from the behaviour with respect to very dilute solutions of basic
dyes that in a number of the objects the gelatin membrane does in fact completely
enclose the cavity filled with gum arabic solution. In these objects the dye is intensely
accumulated in the vacuole and can even bring about coacervation of the gum arabic
inside it (see p. 467, § 5a).

In other objects the "vacuole" remains uncoloured. In this case the complete
envelope has not been established during the preparation or the objects have been
injured during the later treatment and the gum arabic has thus been washed out
of the cavity.

f. The morphology of composite coacervate drops

We stated in Ch. X, § 2 t, p. 378, that at suitably chosen pH and mixing propor-
tions of gelatin (positive), gum arabic (negative) and nucleate (negative), two coexisting
complex coacervates are produced. If one allows them to settle the three liquid layers
arrange themselves according to their specific gravity: at the bottom lies the G + N + a
coacervate layer, then follows G + A + n coacervate layer and on top of that is the
equilibrium liquid. If one shakes the tube, the coacervate layers then divide into
drops which however according to microscopic examination are of a composite
nature[1]: Included in the G + A + n drops there are one or more G + N + a
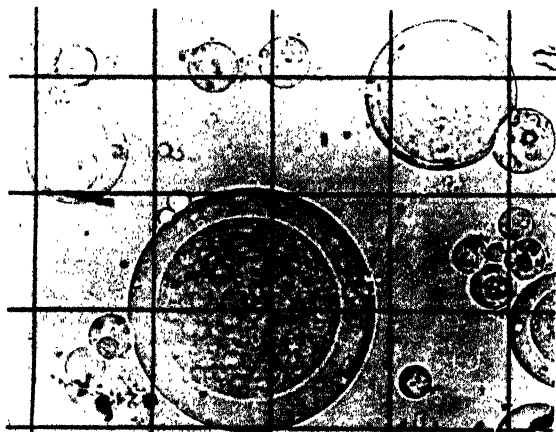
Fig. 5. Composite coacervate drops
consisting of two coexisting com-
plex coacervates (124 × lin.).
Enclosed drops belong to the G +
N + a coacervate (weakly vacuolised),
the surrounding coacervate shells to
the G + A + n coacervate.
The picture was taken at room-
temperature, therefore the coacer-
vates were gelated so that adjacent
drops did not coalesce.

drops (see Fig. 5). Thus the G + N + a coacervate seems to be completely wetted
by the G + A + n coacervate just as any arbitrary organic liquid (p. 436, Fig. 2a).

[1] H. G. BUNGENBERG DE JONG and A. DE HAAN, Bioch. Z. 263 (1933) 33.

The two coacervates possess different affinities for dyes. In the presence of a small amount of neutral salt the coloration of the G + A + n coacervate is practically suppressed, not however that of the G + N + a coacervate. We then see the latter as a coloured sphere lying in the hardly coloured surrounding G + A + n coacervate.

An interesting point is further the position which solid particles absorbed from outside take up in these composite colloid objects [1].

In the example discussed above particles of carbon, calcium oxalate, HgI_2 and MnO_2 are taken up by the G + A + n shell but crowd together at the boundary of the two coexisting coacervates (see Fig. 6 a).
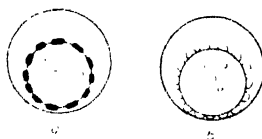
Fig. 6. Position of inclusions (a) or of gelation vacuoles (b) in composite coacervate drops. The inclusions or vacuoles are situated at the boundary of the two coexisting coacervates. The G + N + a coacervate surrounded by the G + A + n coacervate is drawn weakly vacuolised

Obviously these inclusions make a three phase contact with the two coacervates.

g. Vacuoles in three-phase contact

We shall see later on (p. 448) that vacuolation can occur as an accompanying phenomenon in the gelation of complex coacervates containing gelatin. Also the composite coacervate bodies discussed above gelate on cooling and vacuoles may be produced in the peripheral G + A + n coacervate shell. The mechanism of this vacuolation does not however interest us here but only the position which these vacuoles finally take up. This position is the same as that which inclusions take up, that is to say the vacuoles surround the G + N + a coacervate like a garland (see Fig. 6B). This position which is recognisibly the sign of the presence of a three-phase contact, arouses some astonishment.

In fact with vertical observation of the non-gelated composite coacervate bodies we always see the G + N + a coacervate completely surrounded by the G + A + n coacervate and there are thus no indications pointing to the existence of three-phase contact between equilibrium liquid (E) and the two coacervates. Vacuoles are however in principle filled with equilibrium liquid and the position of the vacuoles in Fig. 6 b, which truly points to the presence of a three-phase contact between vacuoles and the two coacervates, is therefore unexpected. The observed position of the vacuoles is open to two interpretations, without it being possible to settle the question of the presence of a three phase contact before gelation:

1. The position of the vacuoles does not represent an equilibrium position belonging to the composite colloid body consisting of two liquid coacervates.

2. The position of the vacuoles does represent such an equilibrium position.

1. Since the enclosed G + N + a coacervate drop gelates somewhat more

[1] H. G. BUNGENBERG DE JONG, Proc. Koninkl. Nederland. Akad. Wetenschap., Amsterdam, 45 (1942) 393.

rapidly than the surrounding G + A + n coacervate shell, shortly before the whole colloid body becomes solid there are present side by side: gelated G + N + a coacervate, liquid G + A + n coacervate and vacuole liquid (E).

It may now be that the wetting proporties of the solid G + N + a coacervate are different from those of the still liquid one and that as a consequence of that the vacuoles come into three-phase contact with them.

2. If we assume that the position of the vacuoles in Fig. 6b really represents an equilibrium position at an instant shortly before the gelation, at which therefore the two coacervates were still liquid, it then follows necessarily from this that the hypothesis tacitly assumed up to now, as regards the complete wetting of the enclosed G + N + a coacervate by the surrounding G + A + n coacervate, is in need of revision. This hypothesis was indeed based on the microscopic picture which these composite colloid bodies exhibit when observed vertically (p. 438, Fig. 5). We are here confronted with the same difficulty which we have mentioned already in the beginning of § 1 e (p. 437) for inclusions: with horizontal observation one observes the picture B of Fig. 3 (p. 437) for our composite coacervate drops.

The G + N + a coacervate has a greater specific gravity than the G + A + n coacervate and it is thus impossible to determine without further information whether the wetting is complete or whether a three-phase contact between G + N + a/ G + A + n/E is present (in which G + A + n is sandwiched between G + N + a and E with a very small angle of contact a).

It is now just the position of the vacuoles in Fig. 6B which might be put forward in favour of the last mentioned possibility; the vacuoles consist of equilibrium liquid E and must then also be attached in three-phase contact to G + N + a, in which G + A + n penetrates with the same very small angle of contact a between the surface of the G + N + a coacervate and the vacuoles V attached to it (see Fig. 7).
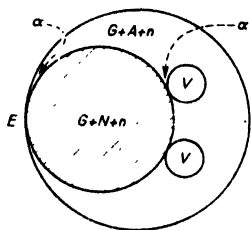
Fig. 7.   Scheme for the discussion of the position of gelation vacuoles in three phase contact. The G + A + n coacervate penetrates with an acute angle of contact a between the vacuoles (small circles) and the G + N + a coacervate (grey). The same acute angle of contact a is also present at the place where the G + N + a coacervate is adjacent to the outer boundary of the G + A + n. coacervate.

The microscopic picture however (Fig. 6b) makes the impression that the gelation vacuoles are attached with a relatively broad basis, but this could be the result of a change in wetting properties of the G + A + n coacervate after solidification of the G + N + a coacervate, as discussed in 1.

h. Significance of the colloid equivalent weight for the morphology of composite coacervate drops

We may ask ourselves why it is just the G + N + a coacervate and not the G + A + n coacervate, which is enveloped completely or possibly almost completely by the second coexisting coacervate.

This must naturally be attributable to the mutual proportions of the three interfacial tensions in question.

$$\sigma_{G+N+a/E} \quad , \quad \sigma_{G+A+n/E} \quad \text{and} \quad \sigma_{G+N+a/G+A+n}$$

If complete enveloping of the G + N + a coacervate by the G + A + n coacervate exists, then we must have:

$$\sigma_{G+N+a/E} \quad > \sigma_{G+A+n/E} \quad + \sigma_{G+N+a/G+A+n} \tag{1}$$

If, as from the evidence of the previous subparagraph we also consider possible, there is just a three-phase contact, but the angle of contact is but little greater than zero, we have approximately:

$$\sigma_{G+N+a/E} \quad = \sigma_{G+A+n/E} \quad + \sigma_{G+N+a/G+A+n} \tag{2}$$

Both cases (1) and (2) have in common that

$$\sigma_{G+N+a/E} \quad > \sigma_{G+A+n/E} \tag{3}$$

and what we must seek is an argument from which (3) follows.

It is now tempting to establish a connection with the results mentioned in § 1 b (p. 433) where we found with the G + A coacervate that the interfacial tension $\sigma_{G+A/E}$ is determined by the intensity of the complex relations.

We can now apply this rule to obtain a verdict concerning the relative magnitude to be expected of the interfacial tensions coacervate/equilibrium liquid in the case of the binary coacervates G + A and G + N. Naturally the other conditions must be comparable (the same pH, the same salt content of the medium, equivalent mixing proportion of G and A or of G and N, that is to say, at their reversal of charge points). Under these comparable conditions the complex relations in the G + N coacervate are considerably greater than in the G + A coacervate on account of the so much smaller equivalent weight of N than of A (see p. 375).

It is thus to be expected that

$$\sigma_{G+N/E} \quad >> \sigma_{G+A/E} \tag{4}$$

We must however obtain a pronouncement on the analogous interfacial tensions in the systems containing all three colloids.

To this end let us consider Fig. 8 (p. 442) which reproduces the upper part of Fig. 33 (see p. 380, Ch. X, § 2 t) magnified. We now consider a total mixture a, situated on the dotted line which joins the two reversal of charge points 3 and 4 on the sides of the triangle GA and GN.

From an investigation [1] into the course of the tie lines in the "ellipse" (the region in which two coexisting coacervates are produced) it has now appeared that this dotted line is itself a tie line so that with the choice of the total mixture a the colloid compositions of these coacervates are given by the intersections 1 (the G + A + n coacervate) and 2 (the G + N + a coacervate) of the dotted line with the ellipse.

Since the points 3, 1, 2 and 4 lie on a straight line, one can consider the coexisting coacervates 1 and 2 as produced by partial mixing of the "binary" coacervates 3 (the equivalent G + A coacervate) and 4 (the equivalent G + N coacervate). In

[1] H. G. BUNGENBERG DE JONG and E. G. HOSKAM, *Proc. Koninkl. Nederland. Akad. Wetenschap.*, Amsterdam, 45 (1942) 387.

consequence of this partial mixing the analogous interfacial tensions coacervate/ equilibrium liquid have taken on other values than they had in the inequality (4) but nevertheless very probably changed in such a way that the larger of the two ($\sigma_{G+N/E}$) becomes smaller by uptake of the third component, the smaller of the two ($\sigma_{G+A/E}$) becomes larger by uptake of the third component.

As a result $\sigma_{G+N+a/E}$ and $\sigma_{G+A+n/E}$ will, it is true, differ less from one another than $\sigma_{G+N/E}$ and $\sigma_{G+A/E}$ but through the fact of the coexistence of the two coacervates this differences cannot become zero.

We thus come to the conclusion that:

$$\sigma_{G+N+a/E} \; \rangle \; \sigma_{G+A+n/E}$$

which we have already deduced above (relation (3)) from the morphological picture of the composite coacervate drops.
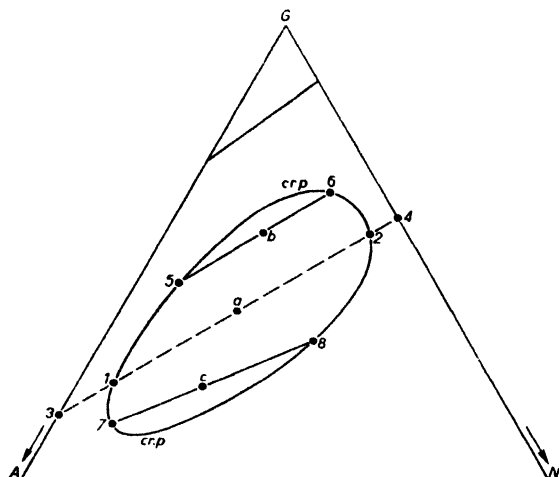
Fig. 8. Situation of the tielines in the region in which coexisting complex coacervates occur. (see text and compare fig. 33, p. 380, Ch. X, § 2 t)
3—4, situation of the electrophoretically uncharged mixtures, at the same time tie line of the uncharged coexisting complex coacervates with the colloid compositions 1 and 2 between the intersections with the ellipse.
5—6 and 7—8 similar tie lines of positively or negatively charged coexisting coacervates of the colloid compositions 5, 6, 7 and 8. Points 1, 5 and 7 are G + A + n coacervates; 2, 6 and 8 are G + N + a coacervates.

The typical morphological pictures of composite coacervate drops discussed up to now are encountered in total mixtures within the ellipse which lie on or near the reversal of charge line (for example a in Fig. 8). The contrast in colloid composition of the coexisting coacervates is the greatest here.

This contrast decreases, in proportion as one chooses as total compositions points within the ellipse which are further removed from the reversal of charge line. This follows from the direction of the tie lines in the ellipse which run approximately parallel to the reversal of charge line. Compare mixtures b or c which give rise to the coexisting coacervates 5 and 6 or 7 and 8 respectively. Since furthermore the absolute values of the interfacial tensions now become smaller (the two coacervates 5 and 6 are positively, the two 7 and 8 negatively charged), the difference in interfacial tension also becomes smaller than with the uncharged coacervates 1 and 2.

It need therefore cause no surprise that with such unfavourably chosen total mixtures the typical picture of Fig. 5 (p. 438) is no longer encountered.

Since $\sigma_{G+N+a/E}$ and $\sigma_{G+A+n/E}$ both become smaller and $\sigma_{G+N+a/G+A+n}$, the interfacial tension between two similarly charged coacervates probably does not change appreciably, the latter can become greater than $\sigma_{G+N+a/E} - \sigma_{G+A+n/E}$, as a result of which the angle of contact, which in the uncharged coacervates is zero or very

small, becomes sufficiently large for its presence to be detected in the microscope image at any rate.

Finally we may mention that composite coacervate drops are also known in which the complex coacervates belong to the type colloid anion ⊢ micro cation[1] (p. 384, Ch. X, § 3). They are produced for example with hexol nitrate in a mixture of sols of:

    a. chondroitin sulphate + arabinate
    b. nucleate + arabinate
    c. nucleate + chondroitin sulphate (the combinations *b.* and *c.* at higher temperatures, since at room temperature nucleate + hexol nitrate does not give coacervation but flocculation, see p. 244, Fig. 6).

The morphological pictures of these composite coacervate drops show that in this case also the equivalent weights of the colloids (see p. 265, Table 1) play a similar part as above in the combination G + A + N.

In the combination *c.* where these equivalent weights differ but little, it does not come to an enveloping of the one coacervate by the other but three-phase contacts are very clearly present.

In the two other combinations where the colloid equivalent weights differ appreciably, morphological pictures can occur (see fig. 9) which show more analogy to the typical pictures in the combination G + A + N. Here also the coacervate, which is rich in the colloid component with the smallest equivalent weight, lies embedded[2] in the coacervate which is rich in the colloid component with the largest equivalent weight.
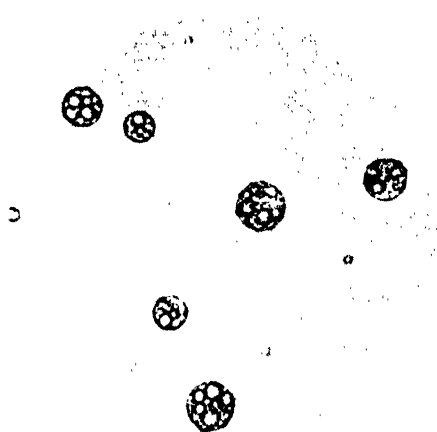
Fig. 9. Composite coacervate drops consisting of two coexisting complex coacervates of the type colloid anion ⊥ micro cation (308 ˅ linear). The drops were produced from a mixture of chondroitin sulphate sol and arabinate sol with hexol nitrate.

§ 2. NON-COMPLICATED VACUOLATION

a. Place, number and size of the vacuoles formed

When through some variable or other the water and micro units content of a coacervate drop increases, one will not in general expect any striking morphological consequences of this[3]. One can on the one hand expect that the coacervate drops

---

[1] H. G. BUNGENBERG DE JONG and A. DE HAAN, *Biochem. Z.*, 263 (1933) 33.
[2] In Fig. 9 the completely or almost completely enclosed chondroitin sulphate-rich coacervate is present in the form of numerous drops.
[3] Coacervates containing not one but two colloids may behave otherwise (p. 472, § 5 b).

become larger through intake of water and micro units. On the other hand however the solubility of the coacervate usually also increases whereby the coacervate drop would become smaller. Which of these two volume changes of the coacervate drops preponderates depends on the volume ratio of the total coacervate and the total equilibrium liquid.

If however through some variable or other the water and micro units content of the coacervate decreases, then striking morphological changes of the coacervate drop are indeed to be expected.
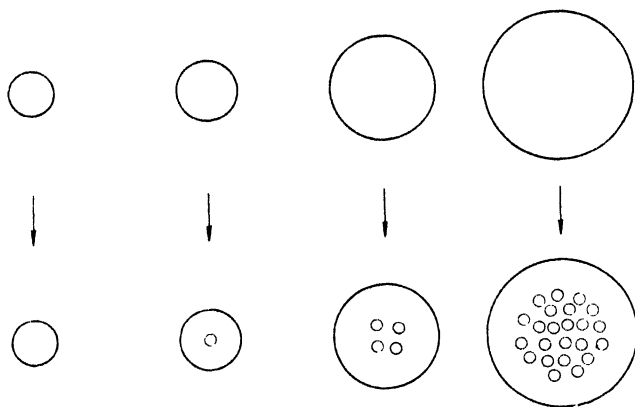
Fig. 10. Dependence of the vacuolation picture of the size of the coacervate drops (see text).

Then there separates out fresh equilibrium liquid in the form of fine drops throughout the whole coacervate drop: vacuolation occurs. It depends on the size of the coacervate drop and the rate at which the variable is changed what character the vacuolation will assume.

If the coacervate drop is very small and the rate not great, the coacervate can lose the excess equilibrium liquid sufficiently rapidly by diffusion to the surroundings so that no vacuolation occurs. With somewhat larger drops one small vacuole is produced in the centre of the drop, with still larger ones several vacuoles (see Fig. 10).

In addition it depends on the rate and the viscosity of the coacervate what vacuolation picture will occur in a coacervate drop of a given size. If the rate is high and if the coacervate is viscous, then the coacervate drops become black by transmitted light, that is to say, more or less opaque through the very large number of very small vacuoles. If the rate is slow and the coacervate not particularly viscous, then the vacuoles first produced can grow in size without fresh vacuoles being

Fig. 11. Coarsely vacuolised coacervate of isoelectric gelatin with resorcinol (108 × lin.). The vacuolation in the original non-vacuolised coacervate drops were produced by slow cooling after preceding heating. On warming a number of drops came into contact with each other and coalesced.

produced again and again. One then obtains a coarsely vacuolised coacervate.

Fig. 11 gives an example which relates to a coacervate of isoelectric gelatin with resorcinol. This coacervate becomes considerably richer in water and resorcinol on rise in temperature. With successive slow cooling coarse vacuolation was produced.

b. Behaviour of vacuoles compared with that of inclusions

Vacuoles, which are produced in a coacervate, are filled with equilibrium liquid. If vacuoles encounter each other accidentally in the coacervate then they coalesce[1]. However even without coalescence the number of vacuoles (leaving gravity out of consideration) will have to diminish in the course of time, since the smaller ones will disappear and the larger ones will increase in volume (compare small drops distilling over to larger ones). In fact the size of the boundary surface equilibrium liquid/coacervate will decrease because of this change and with it the surface free energy will diminish.

Through the difference in sp.gr. from the coacervate vacuoles will however rise[2] and come into contact with the outer boundary of the coacervate drop. These vacuoles then break through and their contents spurt out into the surrounding equilibrium liquid[3]. Here again decrease of the free energy because the total interface coacervate/equilibrium liquid decreases with the withdrawal of vacuoles.

Morphologically there exists a fundamental difference between vacuoles and inclusions such as oil drops, carbon particles etc. These latter are absorbed spontaneously from the equilibrium liquid into the coacervate[4].

This difference is also well brought out in the behaviour of complex coacervate drops in the direct current field. These coacervate drops are in general charged (capillary electric charge at the interface coacervate/equilibrium liquid).

Fig. 12. Behavour of liquid inclusions in complex coacervate drops in a d.c. electric field (schematic). A initial state, B final stage of the relative displacement, in which the organic liquid drop protrudes from the surface of the coacervate drop deformed by the „Büchner effect" (see p. 347). Simultaneous vacuolation phenomena etc. in the complex coacervate drops (p. 347 and 452) are omitted from the scheme.

It now appears that both vacuoles and enclosed organic liquid drops in these coacervates are displaced in the coacervate on the application of an electric field. And indeed they behave as if they have a capillary electric charge on their surface

---

[1] Exception see § 6c, p. 480.

[2] Oleate coacervates have a smaller sp.gr. than the equilibrium liquid. Here the vacuoles will therefore sink in the coacervate.

[3] For an exception see § 6c, p. 480.

[4] From the point of view of the free energy there is no fundamental difference between the exit of vacuoles and the entry of oil drops etc. Since both are spontaneous processes, the free energy decreases in both cases.

of contact of the same sign as the coacervate drop itself. Thus vacuoles and inclusions are displaced in negative coacervate drops towards the anode, in positive towards the cathode. When they reach the periphery of the coacervate drop the vacuoles pass out from it (see p. 348, Fig. 8), the included organic liquid drop does not however. The latter merely protrudes and on removal of the electric field moves back again into the coacervate drop[1] (see p. 445, Fig. 12).

The G + N + a coacervate drops which are surrounded by a shell of G + A + n coacervate in composite coacervate drops (see p. 438), also behave like inclusions. If the G + A + n coacervate is electrophoretically negative, then the enclosed coacervate drop is displaced inside the G + A + n coacervate shell in the direction of the anode; if the surrounding coacervate is positively charged, then it is displaced towards the other side. If the surrounding coacervate is at its reversal of charge point then the enclosed coacervate is not displaced.

c. Vacuolation phenomena in oleate coacervates by volatile and non-volatile organic non-electrolytes

The state of oleate coacervates is influenced by all sorts of organic non-electrolytes (for example even by hydrocarbons[2]). As a result they can become richer or poorer in water.

When one adds such substances to an existing coacervated system, vacuolation of the coacervate drops is therefore to be expected in the second case but not in the first case.

On the other hand on removal of the added organic substance the vacuolation produced in the second case will disappear, in the first case the coacervate drops, which have remained homogeneous, will now begin to vacuolise.
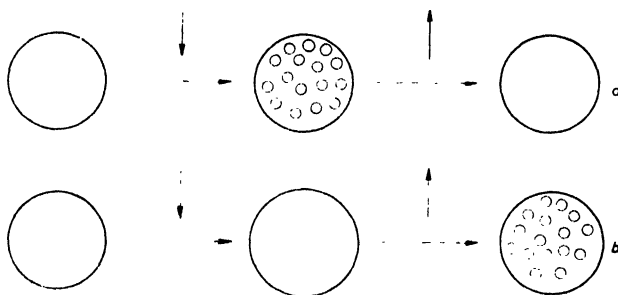
Fig. 13. Vacuolation phenomena with oleate coacervates, in which the agent is added (↓) or removed (↑) via the vapour phase.
a: The agent has a condensing action on the coacervate.
b: The agent has a swelling action on the coacervate.

When the organic substances concerned are sufficiently volatile, one can investigate this with advantage by presenting these substances as vapours to the coacervated system which thus consists of a thin layer of equilibrium liquid with coacervate drops suspended in it.

In this way the organic substance can be added to the system via the vapour phase.

By now exposing the coacervated system afterwards to air without vapour one can remove the organic substance again and thus the possibility of investigating the reversibility of the changes brought about is realised experimentally.

---

[1] H. G. BUNGENBERG DE JONG and A. J. W. KAAS, *Bioch. Z.*, 232 (1931) 338.
[2] Phosphatide coacervates share this property with oleate coacervates (see p. 701, Ch. XIV, § 4).

Such an investigation with oleate coacervates has brought to light the fact that not only the two possibilities mentioned (Fig. 13a and b) but also a third case, in which the first amounts of substance introduced result in vacuolation but a larger amount causes the vacuolation to disappear again. The whole cycle of addition and removal of vapour is then represented by Fig. 14. We shall not deal here with cycles consisting of still more steps.
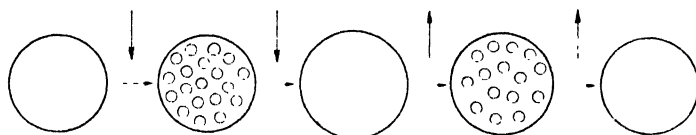
Fig. 14. Vacuolation phenomena in oleate coacervates, in which the agent is added (↓) or removed (↑) via the vapour phase.
The agent produces condensation in small concentrations and swelling in larger concentrations in the coacervate.

Reference may be made to the original publication[1] for further details, the results obtained and biological perspectives. The method described can thus serve to investigate the relation between composition of the organic compound and action on the oleate coacervate. It is however only applicable if the organic substance is sufficiently volatile.

Another method[2] is more generally applicable, in which the morphological consequences of introducing the organic substance into the coacervate itself (isolated from the equilibrium liquid) are investigated (see p. 701, Ch. XIV, § 4).

d. Vacuolation in the gelation of coacervate drops

Coacervates, which contain gelatin, in general gelate on cooling. Exceptions occur when the substance added for coacervation lowers the gelation temperature below room temperature. This is for example the case in the coacervation of iso-electric gelatin with fairly large concentrations of resorcinol. On cooling such coacervate drops only strong vacuolation is produced since the coacervate equilibrium here shifts in the direction of a colloid-richer coacervate on cooling (p. 444, Fig. 11).

Sulphates on the other hand have the effect of raising the gelation temperature of gelatin. One can thus succeed in obtaining gelated coacervate drops by cooling. However these are also vacuolised, since the temperature here shifts the coacervate equilibrium in the same direction as mentioned above. On cooling the gelation does not occur instantaneously and there is thus still some time available to approach to a new coacervate equilibrium before the coacervate solidifies to a gel.

Now it is possible to obtain non-vacuolised gelated coacervate drops by pouring the system coacervated at 50° with $(NH_4)_2SO_4$ into a ten-fold volume of a cold less concentrated $(NH_4)_2SO_4$ solution. The gel globules so obtained show contraction

[1] H. G. Bungenberg de Jong and G. G. P. Saubert, *Protoplasma*, 38 (1937) 329.
[2] H. G. Bungenberg de Jong and G. G. P. Saubert, ibid., 38 (1937) 498.

under the action of tannin (even very dilute solutions of about 0.01% show the effect)
as a result of which they obtain a very highly birefringent membrane[1] (fig. 15A and B).
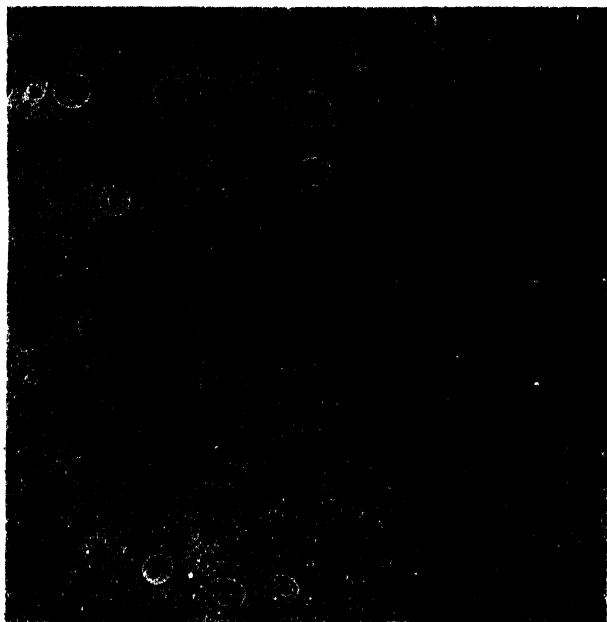
Fig. 15. Action of a dilute tannin solution on gelatin gel globules (40 × lin.).
A: by transmitted light and deep focus. The peripheral condensation membrane is interrupted
at the place where the gel globules lie on the glass surface with a circular plane of contact.

Now it is remarkable that gelatin gel globules obtained in a different way (for
example by emulsification of a concentrated gelatin solution in a liquid not miscible
with water and cooling) do not show this double refraction with tannin or do so
only to a slight extent. It gives the impression that the previous history according
to which the object is obtained (here pseudomorphosis to a coacervate drop) is of
great importance. The question may be asked whether the surface layers of the
coacervate already contain gelatin molecules, oriented in some special way, which
after "compacting" with tannin give rise to the high birefringence.

We now turn to the complex coacervate gelatin — gum arabic. Vacuolation
occurs in this coacervate on cooling also but since the water content is independent
of the temperature at higher temperatures (33—50°), this vacuolation (which occurs
at about 28.5°) is a phenomenon accompanying the gelation itself[2] (and not as in

[1] See also for other properties, H. G. BUNGENBERG DE JONG, Proc. Koninkl. Nederland. Akad.
Wetenschap., Amsterdam, 41 (1938) 646.
[2] H. G. BUNGENBERG DE JONG, E. G. HOSKAM and L. H. V. D. BRANDHOF-SCHAEGEN, Proc.
Koninkl. Nederland. Akad. Wetenschap., Amsterdam 44 (1941) 1104.

Fig. 15.   Action of a dilute tannin solution on gelatin gel globules (40 × lin.).
B: between crossed nicols. →
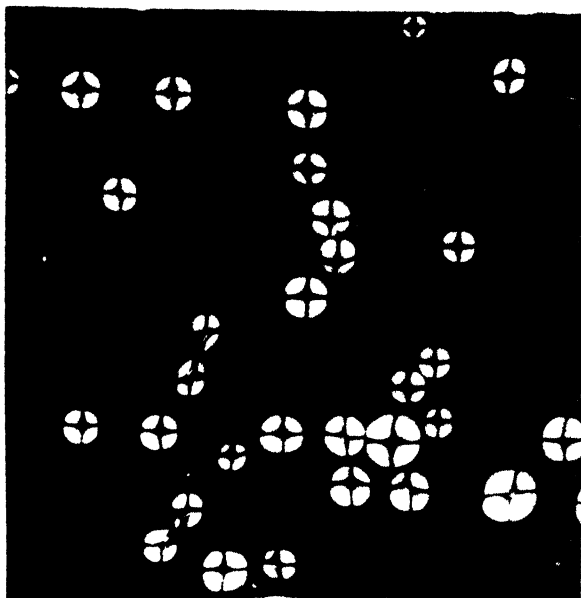
Fig. 16.   Gelatinised complex coacervate drops (gelatin-gum arabic) with coarse vacuolation (after very slow cooling of the drops kept in suspension in their equilibrium liquid) 177 × lin. For the origin of small granular flakes (see p. 382).   ↓
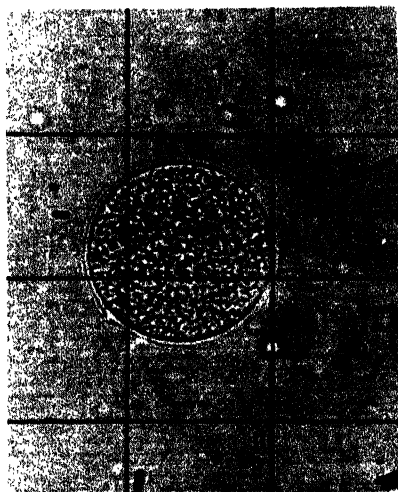
Fig. 17.   Gelatinised complex coacervate drops (gelatin-gum arabic) with very fine vacuolation (after pouring into cold distilled water) 152 × lin.
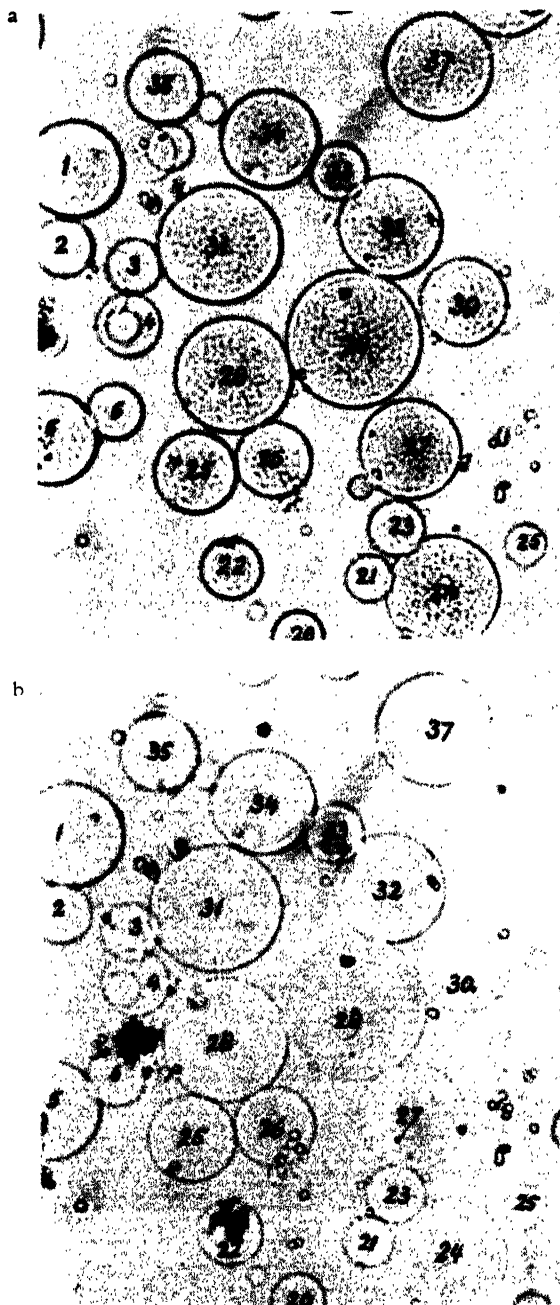
the above (p. 447) mentioned examples mainly a consequence of a displacement of the coacervation equilibrium with the temperature before gelation has occurred).

As far as the morphology of this vacuolation is concerned great variation is again possible[1]. After very slow cooling the gelated coacervate drops contained a number of vacuoles (see p. 449, Fig. 16), after rapid cooling much more smaller ones. In this case the same holds again as was said regarding tempo etc. in § 2a (p. 443).

Gelated coacervate drops can also be obtained by pouring the coacervated system into a large amount (10—100-fold) of cold distilled water.

The very fine vacuolation of these objects, see p. 449, Fig. 17, which appears as fine grains under the microscope is in this case not only to be attributed to the rapid cooling but for a considerable part also to the direct contact of the coacervate drops with the distilled water. As a result of this the salt produced in the coacervation from the counter ions (see p. 368) diffuses out of the drops, and this also gives rise to vacuolation (added salts increase the water content, see p. 364, therefore removal of salts from the coacervate will diminish its water content).

Fig. 18. Reduction of the visibility of the structure of the gelatinised complex coacervate drops by an added salt (see text).

A: original state.
B: after addition of a dilute salt solution. The gel bodies have increased somewhat in volume compared with A. In dropping in the salt solution some gel globules have been washed away.

[1] H. G. BUNGENBERG DE JONG and O. BANK, *Protoplasma*, 33 (1939) 321.

The "complex gels" obtained in the manner described above still show the typical properties of complex systems. They swell with neutral salts (see in more detail p. 381, Ch. X, § 2 u) whereby the optical contrast between gel and vacuoles becomes smaller (compare Fig. 18). At higher salt concentrations the vacuoles become quite invisible and the then greatly swollen gel bodies assume a completely hyaline (transparent) appearance.

§ 3. MOTORY PHENOMENA

a. Motory phenomena in a diffusion field

With suitable coacervate objects one may observe that currents[1] occur in coacervate drops resting on a starched microscope slide (p. 435) when a sufficiently strong diffusion field is produced, in the equilibrium liquid, of a substance which makes the coacervate richer in colloid ("condensing" action). It then appears that the surface of the drop moves towards the side where this condensation (with possibly local formation of small vacuoles as a consequence) is the most pronounced. As a consequence of this the whole contents of the coacervate drop are set in motion as can be seen from the motion of the vacuoles transported with it (Fig. 19a).

When however the diffusing substance makes the coacervate poorer in colloid (swelling action) the motion is reversed (Fig. 19b). In this case the flow is readily visible when one ensures that the coacervate is uniformly provided with small vacuoles by a previous condensation.

The disappearance of the vacuoles (the coacervate becomes hyaline there) then marks the place where the diffusing "swelling" substance strikes the coacervate drop while the direction of the flow is indicated by the motion of the remaining vacuoles. If the coacervate drop lies on a surface which it wets hardly or not at all, then creeping movements can be caused by this movement of the surface and indeed in the case of a condensing substance away from this substance (Fig. 19a to the left), in the case of a swelling substance towards this substance (Fig. 19b to the right).

The above mentioned phenomena can be explained if we assume that condensation of the coacervate is associated with increase of the interfacial tension coacervate / medium, conversely that "swelling" reduces this interfacial tension. (Compare p. 433,

Fig. 19. Streaming in a coacervate drop and direction of creep in a diffusion field (see text).

§ 1 b). Then indeed the interface must be set in motion from places with lower to places with higher interfacial tension.

The assumption made could not so far be tested experimentally in view of the experimental difficulties associated with the measurements of the very small

[1] H. G. BUNGENBERG DE JONG and E. G. HOSKAM, *Proc. Koninkl. Nederland. Akad. Wetenschap.*, Amsterdam, 44 (1941) 1099.

values of σ of the boundary surface coacervate/equilibrium liquid (see p. 435).

For the rest the assumption made is in every way plausible since condensation is associated with separation of fresh small coacervate drops from the equilibrium liquid, that is to say with a reduction of the mutual solubility of coacervate and equilibrium liquid, in the other hand swelling goes with an increase of the mutual solubility. One can now generally expect that in the first case the interfacial tension will increase, in the second case it will decrease.

b. Motory phenomena in coacervates of the dicomplex type in a d.c. electric field

The phenomena which are exhibited by complex coacervate drops in an electric field (type dicomplex coacervates p. 345, Ch. X, § 2e) furnish a fairly complicated case for study in dynamic colloid morphology (p. 17). Here we encounter:

a. Displacement with respect to the medium (electrophoresis),
b. Changes of shape (flattening to discs which are perpendicular to the direction of the field; Büchner effect),
c. Vacuolation and flow phenomena in the interior of the coacervate drops,
d. Transport of large vacuoles to one side of the coacervate drop and breaking through of these vacuoles when they have arrived at the coacervate surface,
e. Formation of small fresh coacervate drops in the medium on the other side of the coacervate drop at some distance from the surface.

Of the effects summarised a, c, d and e are typically polar; they reverse on

Positive   Uncharged   Negative
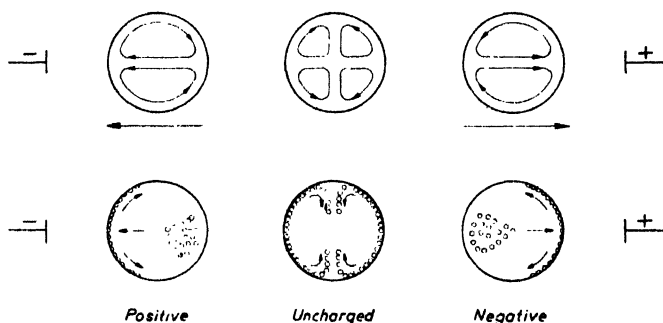
Fig. 20. Streaming and vacuolation phenomena in a complex coacervate drop in a d.c. electric field as depending on the sign of the electrophoretic charge of the coacervate.
Upper row: streaming phenomena and direction of creep movements (horizontal arrows).
Lower row: vacuolation a short time after applying the field.
(In the above given schemes we have deliberately left out of consideration the Büchner deformation which occurs simultaneously).

reversing the polarity of the electric field and furthermore are mirror image with positively charged drops of those with negatively charged ones.

In Chapter X we have already discussed electrophoresis (a), the Büchner effect (b) and points (d) and (e). The phenomena c precede in time somewhat d and e, and as c is of great interest for the interpretation of the "disintegration phenomena"

(that is to say $c + d + e$), they can be best studied in weak fields, which slow down the appearance of d and e.

The flow phenomena[1] occurring therein can be clearly observed even in small fields (5 to 10 volt/cm.) in the complex coacervate gelatin (positive) — gum arabic (negative). In this case a flow occurs with positive coacervate drops in the peripheral layers towards the side of the anode, in negative drops exactly the opposite[2] (Fig. 20 upper row).

As a result of this the interior of the coacervate drop is set in motion in the same way as we have described in § 3a (p. 451) for a drop in the diffusion field of a condensing or swelling substance.

A further point of similarity is the production of vacuoles at the place where the peripheral movement bends inwards[3] (see the triangular vacuolation fields in Fig. 20 lower row).

With very weakly positively charged, uncharged or very weakly negatively charged coacervate drops the flow phenomena do not almost cease but show a more complicated type. In the left half the flow picture then occurs as in positive drops, in the right half that of negative drops (Fig. 20).

It is obvious to seek the cause of the motory phenomena in changes of the interfacial tension coacervate/medium as a result of the polarisation.

In the original publication it was assumed that as a consequence of the field the pH assumes a different value left and right in the drops and this changes the σ values in such a way as to produce the flows described. In the mean time a new investigation[4] has shown that in the buffered system pH changes are not in the first place concerned in the explanation but that the ratio gelatin/gum arabic changes locally.

Since now it has appeared from an investigation[5] on the interfacial tension coacervate/equilibrium liquid that at constant pH this interfacial tension is a maximum at the optimum mixing ratio (that is to say with the uncharged coacervate) but decreases both with excess of gelatin (positive coacervates) and with excess of gum arabic (negative coacervates), (see p. 434, Fig. 1) the elements for an explanation[4] of the vacuolation and flow phenomena in positive and negative coacervate drops are given hereby. Let us take for example, a positive coacervate drop. This is richer in water than the coacervate at the optimum mixing ratio. The gum arabic in the coacervate drop moves in the electric field towards the anode as a result of which

---

[1] H. G. BUNGENBERG DE JONG and E. G. HOSKAM, *Proc. Koninkl. Nederland. Akad. Wetenschap.*, Amsterdam, 44 (1941) 1099.

[2] Compare the observations of FRUMKIN, *J. Colloid Science*, 1 (1946), who describes completely similar movements in an electric field in the case of mercury drops in water.

[3] An independent second system of usually very small vacuoles is however produced on the opposite side, immediately below the interface coacervate/medium (see Fig. 20 lower row). These vacuoles are very closely packed in a single layer and move as soon as they are formed in the direction of the poles of the drop.

This point of detail is missing in coacervate drops in a diffusion field and since its cause is not yet cleared up, this second system of vacuoles is left out of the discussion here.

[4] L. DE RUITER and H. G. BUNGENBERG DE JONG, *Proc. Koninkl. Nederland. Akad. Wetenschap.*, Amsterdam, 50 (1947) 1189.

Further details of the flow phenomena are discussed in this publication, which are not dealt with in Fig. 20 and which are connected with the simultaneously occurring Büchner deformation.

[5] L. DE RUITER and H. G. BUNGENBERG DE JONG, *Proc. Koninkl. Nederland. Akad. Wetenschap.*, Amsterdam, 50 (1947) 836.

the ratio gelatin/gum arabic changes at this side in the direction of the optimum coacervate. Therefore condensation occurs here resulting in vacuolation and increase of the interfacial tension.

At the cathode side of the coacervate drop accumulation of gelatin occurs resulting in a "swelling" action. Thus the coacervate here remains clear while on this side the interfacial tension then becomes lower than originally. As a result of the difference in σ now occurring (on the right higher than originally, on the left lower than originally) the motory phenomena are produced as depicted in Fig. 20 (p. 452). Compare also Fig. 21 I.

The vacuolation and flow phenomena in the negative coacervate drops can be explained in a similar way by again proceeding from the movement of gum arabic towards the anode and of gelatin towards the cathode and alternations in consequences of this in the mixing ratio of the colloids to the right and left in the drop.

From Fig. 21 III it follows from the changes of the original interfacial tension that the flow phenomena in these negative coacervate drops must be just the opposite of those in the positive drops.
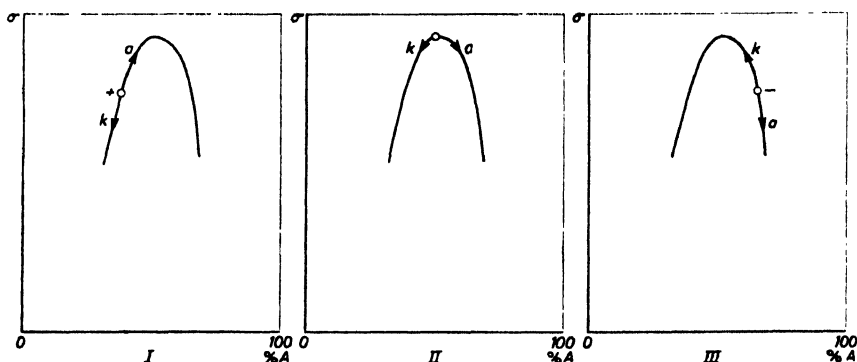
Fig. 21.   Scheme for the explanation of the streaming phenomena in a complex coacervate drop in a d. c. electric field (see text).
I coacervate, positive, II coacervate uncharged, III coacervate negative.
The original interfacial tension is indicated by a point on the curve A of Fig. 1. Arrows indicate in what direction σ changes on the two opposite sides of the drop as a result of a change of the colloid ratio, a at the anode side, k at the cathode side.

The more complicated picture in the almost uncharged drops is also accessible to interpretation. Here σ is originally a maximum and any change left or right can only give rise to a lowering of σ (see Fig. 21 II). As a result of this a streaming of the interface is produced towards the "poles" of the drops (where the field has no oppertunity of producing changes in composition and therefore σ retains its original value). Reference may be made to the original publications for further morphological details and their explanation.

§ 4.   ABNORMAL VACUOLATION

a.   Pulsating vacuoles

In this section a number of vacuolation phenomena are dealt with, in which an abnormal behaviour of the vacuoles in one or more points occurs. While ordinary vacuolation (p. 443, § 2) can be explained in principle by changes in the composition

of the coacervate, we have the impression that in abnormal vacuolation mechanisms are also involved about which we have little information. What is dealt with in this section will be directed rather to a recording of phenomena than to a satisfactory explanation.

If one puts near to one another on a microscope slide a drop of gum arabic solution 2% and a drop of fairly concentrated toluidine blue solution and covers them with a cover glass, then coacervation occurs in the zone of contact of the two liquids. In this zone of contact a gradient of mixing ratios is automatically established, running from exclusively gum arabic sol to exclusively dye solution. Coacervation is an optimum at a certain mixing ratio and the largest coacervate drops form here which are for that matter opaque on account of their large toluidine blue content. The coacervate drops settle out on to the microscope slide and when now either by suction with filter paper or by tilting these drops are surrounded either by the dye solution or by the gum arabic sol they again pass into solution.

One would now expect that in this dissolving of the coacervate the drops become smaller through their peripheral layers passing into solution one by one. This is not however the case when gum arabic sol surrounds the coacervate drops[1].

The coacervate drop gets a central vacuole which becomes larger by absorption of liquid from the surrounding medium. The phenomenon can best be observed when the zone of contact of the two liquids is slowly displaced in the specified direction.

One then sees that the originally opaque drops increase in volume and thereby become transparent at the centre (see Fig. 22).

They now consist of a spherical coacervate shell which is still always strongly coloured but which soon becomes fairly thin and encloses a centrally situated vacuole increasing in volume.
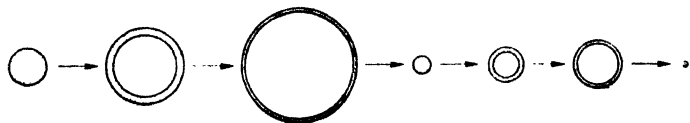
Fig. 22. Pulsating vacuoles in the dicomplex coacervate produced from gum arabic-toluidine blue (see text).

At a certain instant the coacervate shell bursts and changes into a coacervate drop which is smaller than the original one. This can again swell up with the formation of a vacuole and then burst once more. As a rule there is then not enough coacervate over for the whole phenomenon to repeat itself so that the pulsating system is already ruined after a few pulsations (compare also Fig. 43 c, p. 476).

Pulsating vacuoles are also observed in the complex coacervate gelatin — gum arabic if small drops are subjected to a d.c. electric field[2]. In this case also the drop is smaller than the original one after a pulsation.

b. Formation of relatively thick-walled hollow spheres in complex coacervates containing nucleic acid

In general abnormal vacuolation phenomena can more or less be considered as variants on one theme, for which it is a characteristic that one central vacuole

[1] H. G. BUNGENBERG DE JONG, *Proc. Koninkl. Nederland. Akad. Wetenschap.*, Amsterdam, 41 (1938) 643.

[2] H. G. BUNGENBERG DE JONG, O. BANK and E. G. HOSKAM, *Protoplasma*, 34 (1940) 30.

is produced either directly or by coalescence of several small vacuoles but in which this vacuole does not, or does not at first, break through the outer boundary of the coacervate drop. On the contrary the morphological pictures give strongly the impression that at the beginning a mechanism is in action which centres the central vacuole so that the enveloping spherical coacervate shell appears to be of the same thickness everywhere.

This can readily be seen, on account of the relatively large thickness of the walls in these objects, with the hollow spheres formed from drops of the gelatin (positive)-nucleic acid (negative) complex coacervate when indifferent salts are added to the medium[1]. Fig. 23d.

Here we encounter again a similar case to that above with gum arabic + toluidine blue since here also under conditions in which the coacervate will go wholly or partially into solution (in the former case by a change of the mixing ratio colloid anion / dye cation, in the present case by the action of the added salt) a large central vacuole is formed. A point of difference is merely that the vacuole does not break through in the present case and thus no pulsation phenomena occur, at least when one chooses the salt concentration so that the complex coacervate just does not go entirely into solution.
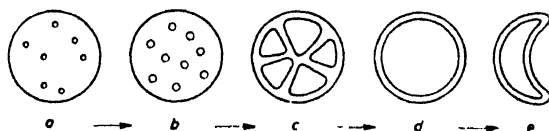
Fig. 23. Influence of added salts on gelatin-nucleic acid complex coacervate drops.
a) originally weakly vacuolised initial state.
b) primary vacuolation after adding salts.
c) and d) subsequent changes.
e) invagination of the gelatinised objects obtained after cooling of d) after addition of saccharose (for invagination processes see p. 464, § 4e).

The central vacuole is here produced (see Fig. 23b—d) by the coalescence of a number of primarily formed smaller ones, whereby the abnormal feature does not lie in the vacuolation itself[2] but rather in the succeeding enlargement of the primary vacuoles and the peculiar contradiction which exists between the coalescence of these vacuoles among themselves and the absence of withdrawal of these vacuoles, or of the central vacuole formed from them, from the drop. This does not occur even after a fairly long time so that the pulsation phenomenon, which we discussed in § 4a, (p. 454) above, is absent here.

In composite coacervate drops (p. 438, § 1f) consisting of coexisting G + A + n and G + N + a coacervates, the latter coacervate, although it now contains a little

[1] H. G. BUNGENBERG DE JONG and C. VAN DER MEER, *Proc. Koninkl. Nederland. Akad. Wetenschap.*, Amsterdam, 45 (1942) 498.
[2] See p. 473: vacuolation of complex coacervates by addition of salt as a result of a change of the colloid ratio in the coacervate.

gum arabic also as well as G and N has retained the property of forming thick walled hollow spheres with salts, at any rate with some salts[1].

Various cases appear thereby according to the salt used (see Fig. 24).
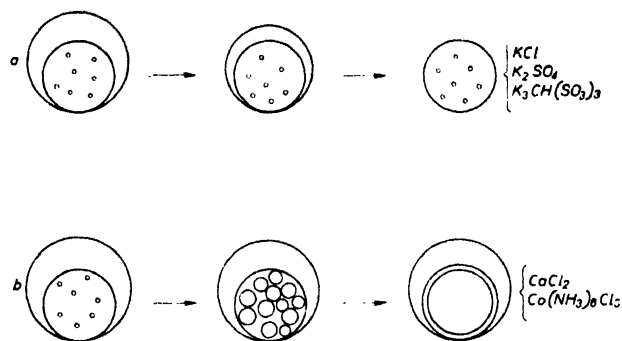
Fig. 24. Influence of added salts on negatively charged composite coacervate drops. The enclosed G + N + a drops were originally weakly vacuolised before the addition of salt. The salts added in a have an influence especially on the surrounding G + A + n coacervate. The added salts in b on the other hand lead to strong vacuolation of the G + N + a coacervate and transformation into hollow spheres.

In Fig. 24a the G + A + n shell is already abolished before the remaining G + N + a coacervate drop is transformed into a hollow sphere. In Fig. 24 b on the other hand the enclosed G + N + a coacervate drop can transform into a hollow sphere while retaining the peripheral coacervate shell. The differences between the salts used must be associated with displacements in the material composition of the coexisting coacervates, in fact we know that salts do this (p. 365) and that the salts arrange themselves in this case in the sequence of the so-called continuous valency rule (p. 452-455).

It is thus to be expected that salts which are situated on opposite sides of this series of salts (for example $Co(NH_3)_6Cl_3$ which is type 3 — 1 and $K_3CH(SO_2)_3$ which is type 1 — 3) will shift the central field in Fig. 33 (p 380) in opposite ways (3 — 1 away from and 1 — 3 towards summit G). The opposite behaviour of the salts in Fig. 24a (types 1 — 3, 1 — 2 and 1 —1) and in Fig. 24b (types 2 — 1 and 3 — 1) must be attributed to this.

The enclosed G + N + a drop in the composite coacervate drops can also transform into a hollow sphere in still another way than addition of salt[2].

This is, among other things, observed by addition of much distilled water in which case the enclosed vacuole withdraws after a long time and thus breaks through into the surrounding G + A + n coacervate. Afterwards the G + N + a hollow sphere rounds itself again to a massive coacervate drop, which lies in the enveloping G + A + n coacervate beside the vacuole.

[1] H. G. BUNGENBERG DE JONG, Proc. Koninkl. Nederland. Akad. Wetenschap., Amsterdam, 45 (1942) 393.
[2] H. G. BUNGENBERG DE JONG and C. VAN DER MEER, Proc. Koninkl. Nederland. Akad. Wetenschap., Amsterdam, 45 (1942) 498.

See Fig. 25 in which the path followed morphologically is indicated by a→c→d→f [1].
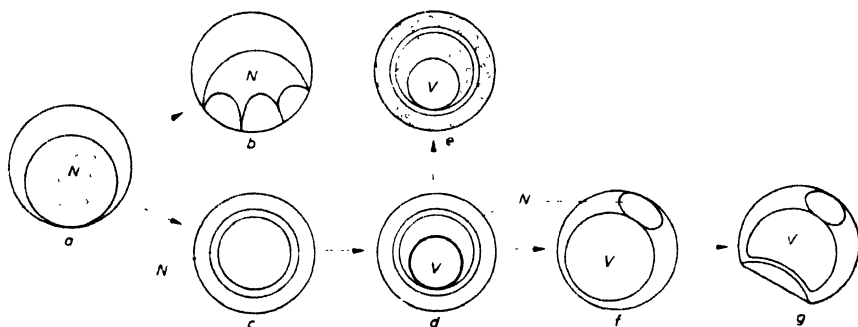
Fig. 25.   Changes of state in the composite coacervate drops on addition of much distilled water.
a)   original state in which the G + N + A coacervate (indicated throughout in the figure by N) was weakly vacuolised. a, c, d and f reproduce changes discussed in the text.
e)   vacuolation occurring when the pH is lowered in d).
g)   invagination of the gelatinised objects obtained after cooling of f) after addition of saccharose (for invagination processes see p. 464, § 4e).

If we compare this with the abnormal vacuolation of the gum arabic — toluidine blue coacervate discussed above (p. 454, § 4a) one can in a certain sense speak of an analogous case: laying down of a large centred vacuole, after a certain time followed by withdrawal of the vacuole. A point of difference is, that there the vacuole bursts out into the surrounding aqueous medium, here however into the enveloping coacervate in which for a certain time it remains embedded.

Further there is another point of difference that in the former case a few pulsations follow each other after which the coacervate has passed completely into solution, in the latter case only a single "pulsation" takes place (the medium conditions are here such that the coacervate also need not disappear).

Yet the most striking difference is that there the withdrawal of the central vacuole is a sudden happening which rightly deserves the name pulsation, here however it is a very slowly evolving process whereby it is possible to follow the path followed morphologically in the withdrawal.

Fig. 25d furnishes just such an intermediate stage, in which the vacuole originaly filling the whole hollow sphere in Fig. 25c has become smaller (V) and between it and the inner wall of the hollow sphere there is a liquid which behaves entirely like the enveloping A + G + n coacervate. (The same type of vacuolation occurs on reduction of the pH for example, in both liquids, see Fig. 25 d→e).

In this stage d the wall of the hollow sphere is obviously already perforated but as a result of the very high viscosity of the G + N + a coacervate this is deformed only very slowly. Naturally it must eventually deform since a perforated spherical shell consisting of liquid cannot be stable from the point of view of the free energy. The morphogical path along which such a perforated coacervate object finally deforms

[1] b is a variant which does not interest us here and f → g is discussed later in the note on p. 465.

into a massive drop is known to us through another object studied[1] (widening of the perforation hole, transition into a beaker shape and further flattening and contraction to the spherical shape).

Nevertheless this deformation is here obviously a sluggish process so that in between the wetting equilibria begin to establish themselves. The inner wall of the G + N + a coacervate shell is wetted better by the G + A + n coacervate than by the vacuole liquid. Therefore this much less viscous coacervate flows inwards through the perforation hole and displaces the vacuole liquid from the wall which is thereby partly pressed outwards (Fig. 25d).

Finally the stage of fig. 25f will naturally be reached by rounding off of the G + N + a coacervate, whereby the now massive G + N + a coacervate lies embedded in a hollow shell of the G + A + n coacervate which surrounds a large central vacuole.

In this fairly long maintained stage it may also be noticed that the vacuole does not or does not at first, break through the G + A + n coacervate to the outside as one would actually expect.

We shall also see below in § 4c that treatment with distilled water can cause similar changes in the G + A coacervate whereby a large enclosed vacuole does not at first break through to the outside.

c. Formation of foam structures and hollow spheres in the gum arabic-gelatin complex coacervate

The primary vacuolation which is produced by the addition of indifferent inorganic salts such as $CaCl_2$, KCl, K_2SO_4 in drops of this complex coacervate, does not in contrast with the gelatin — nucleic acid complex coacervate change into abnormal vacuolation. The vacuoles formed disappear after some time by withdrawal outwards after which the drops are again homogeneous. Striking abnormal vacuolation pheno-
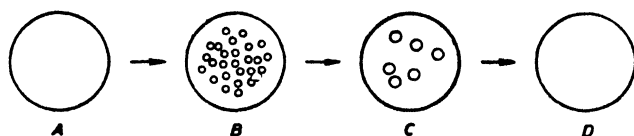
Fig. 26. Vacuolation of the positive complex coacervate drops in contact with much distilled water.
A: original state.
B: primary vacuolation.
C: coarsening of the vacuoles by coalescence.
D: the vacuoles have passed outside.

mena however occur if one brings the coacervate drops into contact with much distilled water at least when the coacervate drops are charged sufficiently negatively[2].

Positive coacervate drops, on the other hand, only show an ordinary primary vacuolation, after which the vacuoles gradually pass outside so that some time later the coacervate drops are again homogeneous. (Fig. 26).

[1] H. G. BUNGENBERG DE JONG, Proc. Koninkl. Nederland. Akad. Wetenschap., Amsterdam, 41 (1938) 3.
[2] H. G. BUNGENBERG DE JONG and O. BANK, Proc. Koninkl. Nederland. Akad. Wetenschap., Amsterdam, 42 (1939) 274.

With negative coacervate drops (see Fig. 27) vacuolation again occurs first of all, followed by a considerable swelling and a mutual flattening of the vacuoles situated directly under the surface of the drop (Fig. 28 a).
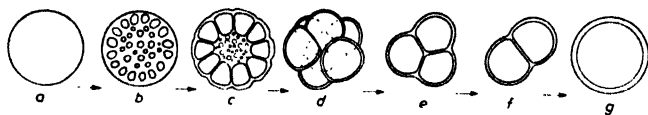
Fig. 27. Morphological development of the hollow spheres from a negatively charged coacervate drop of the complex coacervate gelatin-gum arabic.
a: original state.
b: primary vacuolation.
c and d: foam bodies.
e and f: last stages preceding g.
g: hollow spheres.

This also happens inwards and at a certain stage the coacervate drop is entirely transformed into a mass of "foam vacuoles", the mutually separating coacervate lamellae ;of which gradually burst. It is however remarkable that the lamellae which separate the vacuoles and surrounding medium do not burst so that finally
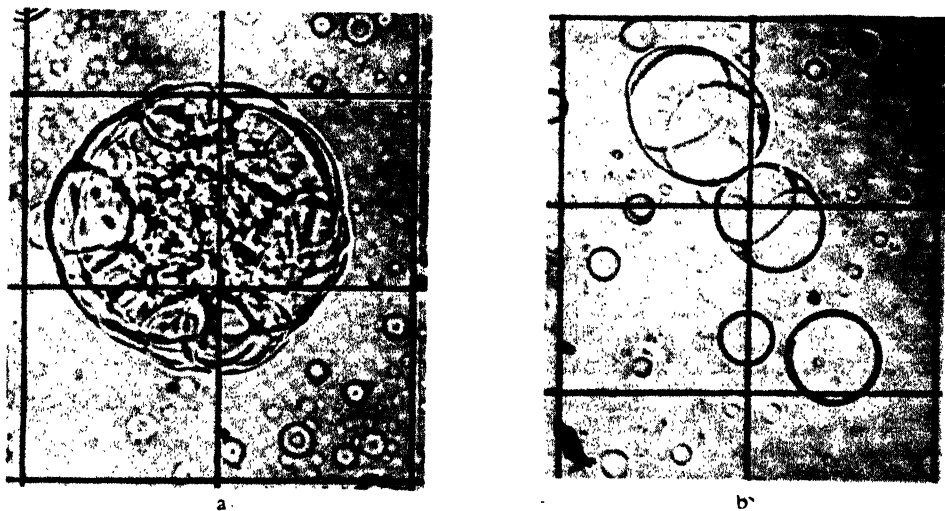
Fig. 28. Foam bodies and hollow spheres produced from negatively charged drops of the complex coacervate gelatin — gum arabic.
A: Foam body corresponding to c of Fig. 27 (203 × lin.).
B: Two foam bodies are still in stage f of Fig. 27, for the rest the coacervate drops have been completely transformed into hollow spheres (196 × lin.).

via stages in which the object consists only of a few foam vacuoles separated by coacervate lamellae it is transformed into a hollow sphere the wall of which can be strikingly thin in typical cases (compare Fig. 28 b).

This typical picture is however not permanent and after some time (for example $^1/_2$ hour) the volume of the central vacuole becomes smaller and consequently the thickness of the coacervate shell greater. Deviations from the first ideal centering of the vacuole now also occur and after a sufficiently long time the vacuole withdraws and one is left with homogeneous coacervate drops.

The primary vacuolation by itself does not particularly present us with a problem of principle. Since there is in the original coacervated system also a little indifferent salt (from the counter ions of the gelatin, for example Cl' and of the gum arabic, for example Ca··, p. 368) a lowering of this salt concentration is produced by the distilled water. The complex coacervate will thus become richer in colloid and consequently the excess of water separates in the form of vacuoles (p. 364. Ch. X, § 21).

An investigation into the composition of the coacervate which remains after withdrawal of the central vacuole confirms various points[1].

This coacervate obtained by "degeneration" of the hollow spheres is indeed poorer in water than the original one. It always contains gum arabic as well as gelatin, the colloid ratio of which is meanwhile somewhat changed in the direction to be expected theoretically from the removal of a salt of the type 2—1 (see p. 365).

Now between the initial state and the final state the abnormal vacuolation process occurs in which the fairly thin walled hollow spheres are produced in a certain intermediate stage. The morphological form of these liquid vesicles then is the expression of two mechanisms acting against one another, the first of which has the upper hand in the formation of foam and vesicles, the second in their degeneration. It is clear that this latter mechanism is caused by the free energy of the macro-interfaces. This free energy always strives for a reduction of the existing surfaces, which can be satisfied by a gradual expulsion of the vacuole content through the intact coacervate lamella of the vesicle (first stages of degeneration) or by expulsion of the complete vacuole to the outside after rupture of the lamella (final stages of degeneration).

There thus remains the question as to what mechanism is concerned in the formation of the foam vacuoles and hollow spheres, which at the beginning stands quite in the foreground. The free energy which is concerned here must be obtained from the original disturbed equilibrium state (bringing the coacervate drops into contact with much water). There are now some indications that in the swelling of the primarily formed vacuoles to foam vacuoles or hollow spheres an electroendosmotic water transport through the coacervate lamellae is involved.

There is agreement with this idea in the fact that small salt concentrations greatly oppose the formation of foam or rapidly bring to degeneration a system already consisting of foam bodies or hollow spheres. $CaCl_2$ in a concentration of a few milliequivalents per 1 makes them transform rapidly into homogeneous coacervate drops. Further the formation of foam bodies and hollow spheres only occurs with sufficiently negatively charged coacervates.

A somewhat more extensive investigation[2] confirms us in this opinion.

---

[1] H. G. BUNGENBERG DE JONG and E. G. HOSKAM, *Proc. Koninkl. Nederland. Akad. Wetenschap.*, Amsterdam, 45 (1942) 200.

[2] H. G. BUNGENBERG DE JONG, O. BANK and E. G. HOSKAM, *Protoplasma*, 34 (1940) 30.

The conditions already discussed are expressed in Fig. 29 from which one can read off at what mixing ratios and pH's of the original sols foam bodies and hollow spheres are formed after shaking the isolated coacervate layer with a certain amount of "isohydric distilled water" (that is to say, dilute HCl of the same pH). This takes place in the shaded strip.

The outermost ellipse A indicates the mixing ratios and pH's at which the original coacervate volume is just zero. The coacervation region therefore lies inside this ellipse. The position of the reversal of charge points of the original coacervates is indicated by the slightly curved thick line a inside the ellipse. This divides the ellipse therefore into two regions, the left one of which represents the positive coacervate drops, the right one the negative. We thus see that the foam region (the shaded bar) lies in the region of the negative coacervate drops.

The dotted ellipse J and the dotted straight line similarly give the coacervate region and the reversal of charge after treatment with "isohydric distilled water". The coacervation region has therefore shrunk, which can be explained because a new equilibrium liquid has formed from the "isohydric distilled water" at the expense of coacervate gone into solution.
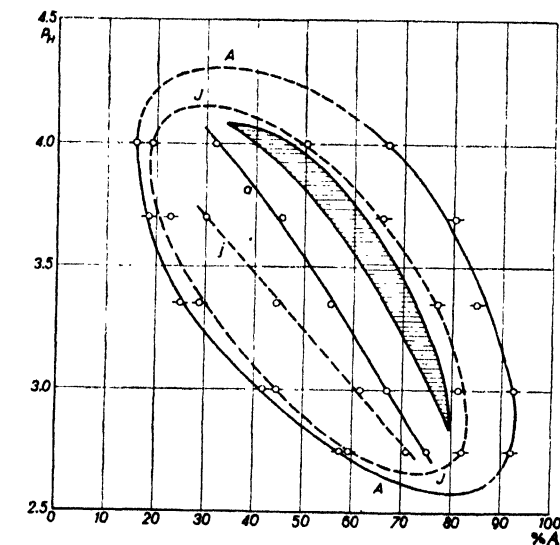
Fig. 29. Figure epitomising a part of the conditions for the formation of hollow spheres (see text.) The coacervate is electrophoretically negative both originally and finally in the shaded region of corresponding pairs of pH values and mixing ratios.

Further the axis of this ellipse J lies more to the left than that of the ellipse A and similarly the reversal of charge line j (dotted one) compared with the fulll line is also displaced to the left. A negativization has thus occurred after the system has been brought into contact with distilled water, which can be attributed to a removal of CaCl$_2$ from the total system (in fact addition of CaCl$_2$ results in positivization in a complex coacervate, see p. 364, Ch. X, § 21). We see from the figure that the region in which formation of foam bodies and hollow spheres takes place lies near the boundary of the ellipse J and indeed on the side where the charge of the coacervate, both before and after foam formation, is negative. It can be assumed from this that the charge of the coacervate is always negative even during the foam formation.

To test the hypothesis that an electroendosmosis takes place in the formation of foam bodies and hollow spheres the E.M.F. of cells of the following type was also determined:

calomel electrode	sat. KCl	"isohydric dist. water"	coacervate	equil. liquid	sat. KCl	calomel electrode.

as a model for the situation presenting itself at the first instant:

"isohydric distilled water" | coacervate | contents of the primary vacuole.

The results (see Fig. 30 in which only the ellipse J and the corresponding electrophoretic reversal of charge line j are included in the figure), show that the shaded region is bounded on its left side by those states of the original coacervate in which the above mentioned E.M.F. is just zero (see the dotted curve). To the right of this the "isohydric water" in the cell is positive with respect to the contents of the vacuole.

We thus come to the conclusion that for the formation of foam bodies and thin walled hollow spheres two conditions must be fulfilled simultaneously:

1. The coacervate lamella must be electrophoretically negatively charged;

2. The vacuole contents must be negative with respect to the "isohydric distilled water". These two conditions speak very much in favour of an electroendosmotic water transport through the negative coacervate lamella to the vacuoles.

Nevertheless we are still far removed from a really satisfactory explanation, for in fact an electroendosmotic water transport to the primary vacuoles ought to come about in the mirror image situation, viz coacervate lamella positive and vacuole contents positive with respect to the "isohydric distilled water".

Fig. 30. Figure epitomising the two conditions for the formation of hollow spheres.
Apart from the shaded region only the ellipse J and the corresponding curve for the electrophoretically uncharged state (curve $\zeta = 0$) have been taken over from Fig. 29. The shaded region is bounded on the left by the curve which represents the zero E.M.F. points of the chains discussed in the text.

This mirror image situation is present in Fig. 30 in the area of the ellipse to the left of the $\zeta = 0$ line.

The primary vacuoles do indeed here also grow with the time but this is transitory and can possibly also be the result of coalescence of smaller vacuoles. We do not however here come as far as the formation of typical foam bodies.

It is therefore not yet possible to give a reason why the production of typical foam bodies is associated with negative coacervates.

d. Properties of the hollow spheres and of the massive coacervate drops which are produced from them by degeneration

The hollow spheres discussed in § 4c behave differently with respect to their wetting behaviour than the original coacervate drops. While the latter spread over

a glass surface and thus can only be viewed permanently with the microscope on starched microscope slides (p. 435), this is not the case for the former. They also do not coalesce mutually when they come into contact with each other. This peculiarity is also shown by the massive coacervate drops which are gradually produced from them by degeneration.

When they settle out by sedimentation they form a close packed arrangement. If one cools them they gelatinize and one can shake them loose as separate gel globules. It gives the impression that these objects have therefore a different (no longer decisively liquid) surface.

The nature of these changes is not yet known. One can however manage to make these massive drops coalesce to a homogeneous liquid layer by applying vigorous centrifuging. The coacervate surface again becomes typically liquid as soon as one adds a little indifferent salt (a few m. eq. p. 1).

As already stated the hollow spheres after they have formed from foam bodies continue to exist for some time (for example $^1/_2$ hour) in their typical form.

During this time therefore the processes, which on the one hand gave rise to their formation, on the other hand threaten them with degeneration, are maintained in equilibrium. The former, the formative process, originally much in the ascendent, has gradually decreased in intensity but not yet so far that the degeneration process proceeds noticeably rapidly. In this quasi-stationary state one can however reintroduce the formative process by applying further negativization.

If the pH of the medium containing the suspended hollow spheres is increased to a slight extent (by carefully adding a little NaOH) foam formation again occurs, now however in the spherical shell of coacervate (Fig. 31 a→b). The original central vacuole can thereby become much smaller and the microscopic image is then completely dominated by the vacuoles newly produced in the coacervate shell and greatly enlarged. After some time the foam structure disappears spontaneously and the typical hollow sphere is restored (Fig. 31 b→c).
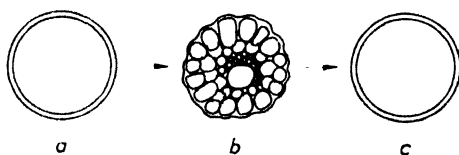
Fig. 31. Transitory foam formation of a hollow sphere after slight increase of the pH.
a) original state.
b) many foam vacuoles have been produced in the coacervate shell of the hollow sphere, whereby the originally large central vacuole has become small.
c) final state produced spontaneously after some time.

e. Gelatinized hollow spheres. Invagination processes

On cooling the hollow spheres gelatinize and the spontaneous degeneration now no longer follows. Invagination processes now occur in these gelatinized objects under all sorts of conditions (Fig. 32)[1]. Among these phenomena are some, the cause of which one is inclined to seek in an osmotic abstraction of water from the vacuole. The invagination is then very deep (for example Fig. 32d). This occurs, for example

---

[1] H. G. BUNGENBERG DE JONG, O. BANK and E. G. HOSKAM, *Protoplasma*, 34 (1940) 30.

when they are placed in solutions of cane sugar. The invagination disappears again rapidly in distilled water. Eventually however the indentation also disappears spontaneously (in cane sugar for example after 1 hour).

The wall of the gelatinized hollow spheres must therefore be fairly difficultly permeable to cane sugar. With organic non electrolytes there is a connection between molecular size and the time after which the indentation disappears spontaneously. This time increases very greatly in the order methyl alcohol — glycol — glycerol — cane sugar[1].

Salts also give rise to invagination; the latter is usually less deep (Fig. 32b or c) and goes back spontaneously very quickly.
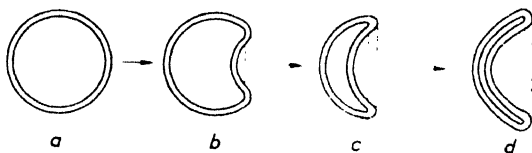
Fig. 32.   Invagination of a gelatinized hollow sphere (see text).

We are inclined here to seek the cause of the invagination not in the first place in an osmotic abstraction of water from the vacuole but, concurrently or possibly mainly, in the action of the salt on the wall of the hollow sphere. Since this consists of a gelatinized complex coacervate and salts have a swelling action on it (see p. 383, Fig. 34) the wall of the hollow sphere, when salt penetrates, must not only increase in thickness but also the whole diameter of the hollow sphere must become larger. This hollow sphere is however filled with vacuole liquid, which resists the volume increase of the vacuole. When the swelling of the wall proceeds more rapidly than the vacuole can manage to add water through the wall from outside, invagination must result. The elastic forces of the deformed wall form the driving force by which water is sucked in from outside and by which the depression shortly rounds itself off again.

Finally there are also circumstances, which lead to invagination, which are such that one can with certainty leave an osmotic abstraction of water from the vacuole out of consideration. This occurs, for example, on adding distilled water to the medium in which the gelatinized hollow spheres are suspended: the invagination produced, usually not very deep, (Fig. 32b) again belongs to the type which rapidly disappears spontaneously but too little data are yet available for an explanation.

f. Colloid displacements in gelatinized hollow spheres towards the vacuole

If the gelatinized hollow spheres are subjected to a certain preliminary treatment (remaining some time in tap water) and then placed in a very dilute solution of neutral red in distilled water, an accumulation of neutral red in the vacuole takes place, while the surrounding spherical shell itself remains almost uncoloured[2]. Fig. 33a→b.

---

[1] The G + A + n shell of the objects depicted in Fig. 25f (p. 458) also give invagination after gelatinization: f ⟶ g which again disappears after addition of distilled water. This is similarly the case with the gelatinized relatively thick walled G + N hollow spheres of Fig. 23 (d ⟶ e), see pag. 456.
[2] H. G. BUNGENBERG DE JONG, O. BANK and E. G. HOSKAM, Protoplasma, 34 (1940) 30.

After the accumulation has reached a certain degree, the very intensely coloured vacuole liquid "demixes"; a number of very small intensely coloured drops separate
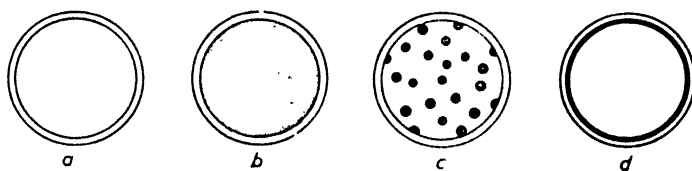
Fig. 33. Accumulation of a basic dye and coacervation in the central vacuole of a gelatinised hollow sphere.

a: original state, by a previous treatment (see text) the vacuole contains arabinate sol.
b: diffuse accumulation of the basic dye.
c: coacervation of the arabinate by the dye cations (dicomplex coacervation colloid anion + micro cation type, Chapter X, § 3d, p. 392).
d: the intensely coloured coacervate has flowed out against the inner wall of the hollow sphere (drawn in optical cross section).

out which are initially in strong Brownian movement. The drops coalesce with each other and deposit on the wall of the vacuole (Fig. 33c), while the remaining vacuole liquid — under microscopic examination — is now colourless (compare Fig. 34). The intensely coloured drops spread gradually over the wall of the vacuole as a result of which the latter becomes covered with a thin layer of coloured liquid (Fig. 33D).

The accumulation and the "demixing" succeeding it are in this case just made possible by the preceeding treatment (pH increase by the tap water which reacts somewhat alkaline). As a result of this the object is subjected to conditions (pH higher than the isoelectric point of the gelatin) in which the complex relations in the original complex gel of the spherical shell are lost and the gum arabic (being in principle a mobile complex component, see p. 383, Ch. X, § 2u) can gradually diffuse out of the wall. The vacuole liquid thereby gradually assumes the character of a gum arabic sol. The accumulation and the coacervation succeeding it do not differ further in anything from similar cases in which

Fig. 34. Accumulation of neutral red in the central vacuole of the hollow sphere, representing stage c of Fig. 33 (160 × lin.). In the left square on the second row from above there is a hollow sphere which has not shown a typical accumulation. The irregular shape suggests that this has been damaged as a result of which the arabinate sol was washed out.

colloid objects contain a vacuole filled with gum arabic sol (see p. 437, § 1e, and p. 474, § 5c).

§ 5. COLLOID SYSTEMS SURROUNDED BY TOTALLY CLOSED MEMBRANES

a. Morphology after coacervation of the enclosed colloidal system

A first attempt[1] to enclose colloid systems in a collodion membrane, in order to be able to study reversibly the behaviour of that colloid system (for example coacervation) when liquids are made to flow past this membrane, already furnished encouraging results.

In this method a small amount of a hydrosol, for example, a gum arabic solution, is coarsely emulsified in a solution of celloidin in a mixture of ether and amyl alcohol and this emulsion is applied in a thin layer on a glass plate. After allowing the ether to evaporate for some time, the film is placed in water. There are then round inclusions of the hydrosol in the membrane. If one makes a very dilute solution of a basic dye (for example toluidine blue) flow past this membrane, a considerable accumulation of the dye occurs in the enclosed gum arabic sol and coacervation after some time. See Fig. 35.

If a mixture of gelatin and gum arabic sols is enclosed in the membrane, then one can see the com-
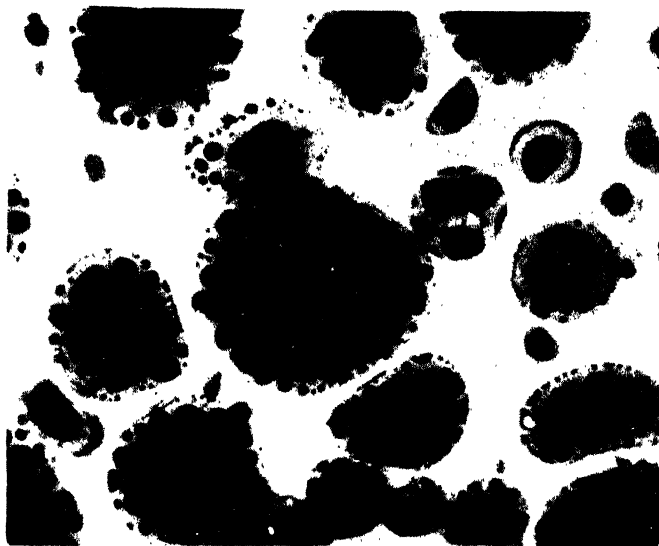
Fig. 35. Coacervation of gum arabic sol enclosed in the cavities of a celloidin membrane; after exposure to a very dilute toluidin blue solution. (200 × lin.).

plex coacervation accomplished when one makes very dilute acetic acid flow past the membrane while it is warmed.

In these investigations it had already been found that in the final state a coacervate wets the walls of the enclosing membrane and thus a parietal coacervate is produced which surrounds a central cavity filled with equilibrium liquid. Because of the comparatively large thickness of the celloidin membranes (diffusion processes slow), it lasts a considerable time before this final state is reached.

The method mentioned has also other disadvantages and frequently furnishes

[1] H. G. BUNGENBERG DE JONG and O. BANK, *Proc. Koninkl. Nederland. Akad. Wetenschap.*, Amsterdam, 42 (1939) 83.
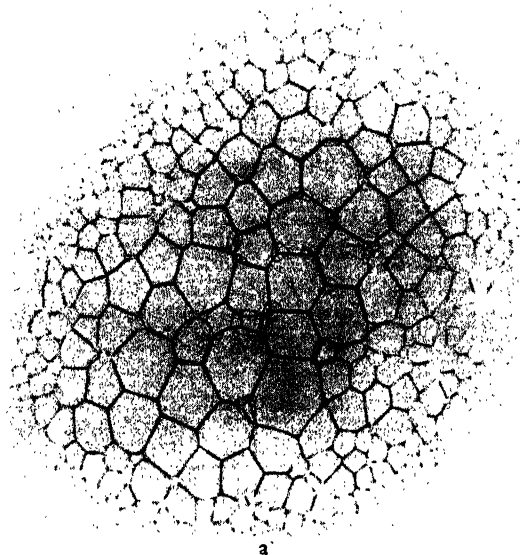
membranes in which the cavities are perforated. In addition they do not lie in one plane but often in layers on top of each other. A great improvement occurs in all respects when the original emulsion is not allowed to dry on glass but is spread on the inter-face water/air[1].

With the correct proportions of amount of emulsion spread and size of the water surface thin membranes are obtained in which there are groups of contiguous prismatic

Fig. 36. Artificial tissue of prismatic cells, as used for the study of biocolloid systems totally enclosed in celloidin membranes.
a: One of the groups of contiguous cells (93 × lin.).
b: Shape of the cell compartments seen in perspective (330 × lin.).

cell compartments (Fig. 36a) in which the original hydrosol is enclosed. Fig. 36b gives an impression of the shape of the cell compartments in the membrane.

One can now lift the membrane from the water surface with a microscope slide and cover an open cuvette with this microscope slide as a lid. By making a medium of known composition now flow past the membrane one can study the influence of it on the enclosed colloid system.

If for example, a mixture of gelatin and gum
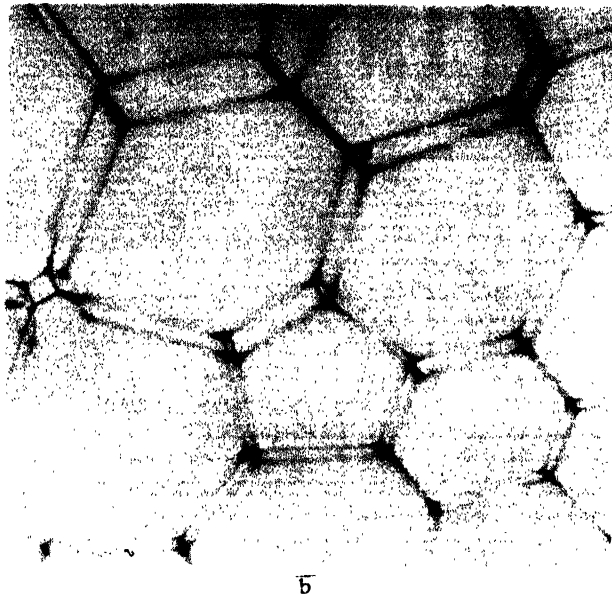
[1] H. G. BUNGENBERG DE JONG, B. KOK and D. R. KREGER, Proc. Koninkl. Nederland. Akad. Wetenschap., Amsterdam, 43 (1940) 512.

arabic sols is enclosed, then after passage through the cuvette of warm (35°) dilute acetic acid (for example of pH 3.5) complex coacervation is produced in the cell compartments. In the final state the complex coacervate again wets the cell walls as a coherent liquid layer and then encloses one central vacuole filled with equilibrium liquid (Fig. 37).

According to the circumstances (p. 470, Fig. 38) this final picture can be reached along three different ways:

a

b
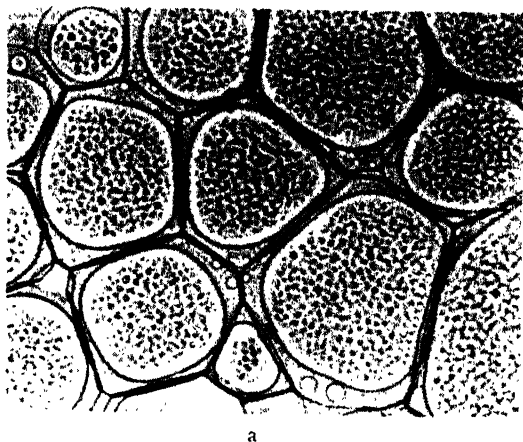
Fig. 37. Complex coacervation of a mixture of gelatin and gum arabic sols enclosed in cell compartments.
a: The complex coacervate has separated out as drops; they coalesce with each other and flow on contact with the wall over it with the formation of a coacervate layer adjacent to the wall.
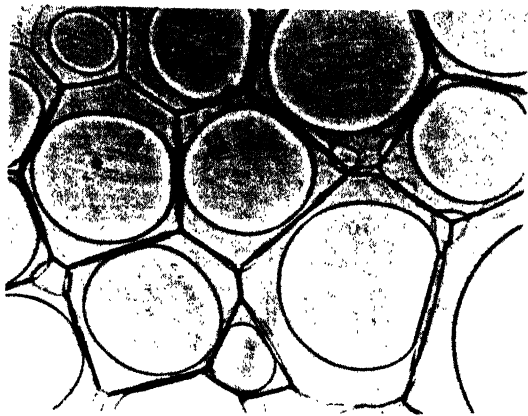
b: Final state, in which all the coacervate drops have fused into a microscopically homogeneous coacervate layer adjacent to the wall and surrounding a central „vacuole".

1. coacervate drops are produced which coalesce with each other (Fig. 37 and p. 470, Fig. 38a),
2. vacuoles are produced which coalesce with each other to form the central vacuole (Fig. 38b),
3. when putting the membrane (of which the contents of the cells have been gelatinized by precooling) on a cold cuvette, then on warming a more or less central lump is produced which is connected to the wall by threads and strands. Gradually the whole becomes more fluid and the usual final picture is produced (Fig. 38c).

A short explanation of the above three ways seems at place. The first way occurs if the volume ratio coacervate/equilibrium liquid in the very beginning of the coacervation is zero and gradually increases.

The second way occurs if the volume ratio coacervate/equilibrium liquid at the first moment of coacervation is infinite. It has e.g. been realised by first adjusting the necessary pH in presence of an indifferent salt, and then diminishing the con-

centration of that salt, without altering the pH. In the second treatment the suppressing action of indifferent salts on complex coacervation (see p. 349, Ch. X, § 2 f) is gradually

removed and at a certain moment the whole compartment is filled with a very waterrich coacervate, which by vacuolisation changes in a coacervate poorer in water (the effective attraction is increased by removal of the indifferent salt, see p. 364).

For the third way one must start with a gelated sol mixture in the cell compartments. This gel changes after adjustment of the pH into a complex gel (p. 381, Ch. X, § 2u), which at sufficiently low temperature needs not yet be accompanied by morphological changes. It softens however by heat and gradually changes into the ordinary two layer system coacervate/equilibrium liquid.

Fig. 38. Different ways along which the final morphological position of the complex coacervate gelatin-gum arabic in the cell compartments can be reached ·(see text).
a: 0.01 N acetic acid is passed through the cuvette at 35° C.
b: 0.01 N acetic acid, containing 40 m. eq. p. 1 NaCl is passed at 35°, followed by 0.01 N acetic acid without salt.
c: cold 0.01 N acetic acid is passed through the cuvette and the temperature is gradually rised.

If a mixture of gelatin, gum arabic and Na nucleate is enclosed, then coexisting complex coacervates are produced in a similar way, whereby finally the G + A + n coacervate, as the coacervate adjacent to the wall, surrounds the central vacuole, while the G + N + a coacervate is enclosed in this parietal coacervate as one (Fig. 39c) or more (Fig. 39 a en b) drops [1].
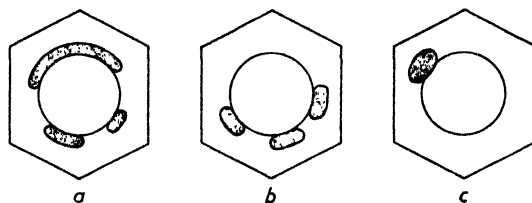
Fig. 39. Topographical position of the three coexisting liquids (after complex coacervation of a gelatin + arabinate + nucleate sol mixture) in the cell compartments of an artifical tissue (schematic). The G + N + a coacervate (grey) lies embedded in the G + A + n coacervate adjacent to the wall. This latter encloses a central vacuole (equilibrium liquid).

It is here again impossible to determine from the morphological picture whether the G + N + a coacervate really lies entirely separated from the vacuole by a very

[1] H. G. BUNGENBERG DE JONG, Proc. Koninkl. Nederland. Akad. Wetenschap., Amsterdam, 45 (1942) 76.

thin G + A + n lamella or whether there is a three phase contact of the G + A ! n coacervate with a very acute angle of contact (this difficulty has already been mentioned in § 1g, p. 439 and § 1h, p. 440).

The method sketched above can be used with advantage in the study of dynamic colloid morphology in which we are set the task of studying the path morphologically followed from one equilibrium state to another as well as investigating the reversibility of this path.

For the rest this method is not usable in every case.

In attempts to enclose native protein sols one encounters the difficulty that denaturation occurs at the surface of contact of sol and emulsifying liquid. Further it is impossible to enclose sols of typical lipophilic association colloids, such as oleates and phosphatides since with these the membranes break on spreading.

The cell walls are permeable to some relatively low molecular colloids so that they are more or less rapidly washed out from the cell compartments. This is for example the case with yeast nucleic acid, as a result of which the composition in a cell containing G + N + A changes continually. As the coalescence of very fine coacervate drops to the states represented in Fig. 39 requires a comparatively long time and as the composition of the content of the cell must lie within narrow limits (c.f. Ch. X, § 2t, Fig. 33, p. 380) the realisation of satisfactory morphological pictures meets with serious difficulties and indeed has only succeeded one or two times. Moreover the states of Fig 39 do not represent the final equilibrium but in the end the coacervate G + N + A disappears completely through continued loss of nucleic acid.

With regard to the formation of groups of contiguous prismatic cells in the spreading (p. 468, Fig. 36) we may further mention that the enclosed hydrosol itself plays no part in the process since the groups are obtained just as well in the emulsification of distilled water. If one shakes water with the emulsifying liquid, then this is divided into small drops which no longer coalesce. From the celloidin organosol a colloid film of nitrocellulose forms on the surface of the water drops.

In the speading of the emulsion on the water/air boundary these drops in the still just liquid layer drift together into groups as a result of which they stick together superficially. On solidification of the membrane they are flattened to a further extent and assume the prismatic shape as a result of which the groups are transformed into tissue patches of cells.

The factors which are of importance for the successful production of these membranes have been studied in more detail[1]. There are various circumstances in which the cell compartments can be injured. It may suffice here to remark that for the integrity of the cell walls a mixture of ether + amyl alcohol + benzene (2 : 1 : 1) has proved more satisfactory as a solvent for the celloidin then the original mixture of ether + amyl alcohol.

---

[1] H. G. BUNGENBERG DE JONG and R. C. BAKHUIZEN VAN DEN BRINK, *Proc. Koninkl. Nederland. Akad. Wetenschap.*, Amsterdam, 51 (1948) 3.

b. Dynamic colloid morphology of the parietal gelatin + gum arabic complex coacervate enclosed in the cell compartments

In § 5a we discussed the final positions which the coacervate and the equilibrium liquid assume with respect to one another and the enclosing membrane. This still belongs to static colloid morphology. One can however with precisely the method described in § 5a also study objects which can be regarded as belonging to dynamic colloid morphology. In this, after the final state (homogeneous parietal coacervate) has been established the liquid flowing along the membrane may, for example, be changed in composition and the morphological path studied along which the colloid system changes until it has again arrived at equilibrium with the new liquid flowing past. Then one again switches over to the original liquid by which one can thus determine whether the changes, which have occurred, are reversible. The complex coacervate gelatin—gum arabic has been studied in detail in this way, the influence of a change of pH[1] on the one hand and of salts or non-electrolytes[2] having been investigated. Further the influence of an electric field on the complex coacervate has also been investigated, in which it appeared that the resultant changes in pH on either side of the celloidin walls separating the cell contents play an important part[3].

Fig. 40. Morphological changes on alteration of the pH of the liquid flowing past the cell membrane (see text).
Middle row: small alterations of pH (0.25 pH) have as yet no morphological consequences.
Upper and lower row: larger pH changes (0.5 pH) produce vacuolation in the complex coacervate adjacent to the wall. This vacuolation on restoring the pH does not decrease immediately but rather increases, while in addition small coacervate drops separate out in the central vacuole.

We shall restrict ourselves here to the influence of the pH and of salts. It appeared that starting from the optimum coacervate (most favourable mixing ratio at the chosen initial pH) both an increase of pH and a decrease (at least if it is not too small) brings about vacuolation in the coacervate adjacent to the wall. See Fig. 40.

This is surprising at first sight, because with the fixed mixing ratio of the two colloids in the compartment any change in the chosen optimum pH necessarily leads to a coacervate with higher water-content. Compare p. 362, Ch. X, § 2k; p. 363, Fig. 23; p. 368. This would thus be a divergence from the rule which was given in § 2a (p. 443): decrease of the water-content leads to vacuolation, increase however does not.

[1] H. G. BUNGENBERG DE JONG and B. KOK, Proc. Koninkl. Nederland. Akad. Wetenschap., Amsterdam, 45, (1942) 51.
[2] H. G. BUNGENBERG DE JONG and B. KOK, Proc. Koninkl. Nederland. Akad. Wetenschap., Amsterdam, 45 (1942) 67; 45 (1942) 204.
[3] H. G. BUNGENBERG DE JONG and D. R. KREGER, Proc. Koninkl. Nederland. Akad. Wetenschap., Amsterdam, 43 (1940) 732.

It should however be borne in mind in this connection that this rule only holds strictly for coacervates containing only one colloid. With complex coacervates the ratio of the two colloids in the coacervate can also change at the same time.

This is indeed the case here. Compare in p. 363, Fig. 24, Ch. X, § 2k, from which it appears that proceeding from pH 3.5 to higher or lower pH also changes the relative proportion of the two colloids in the coacervate.

The vacuolation observed in the coacervate adjacent to the wall on increase or decrease of the pH has as its object the achievment of the colloid ratio corresponding to the new pH in the coacervate. The vacuolation must therefore not be interpreted in this case as a sign that the coacervate becomes poorer in water.

On the contrary, the coacervate must in fact take up water simultaneously with this vacuol-

a
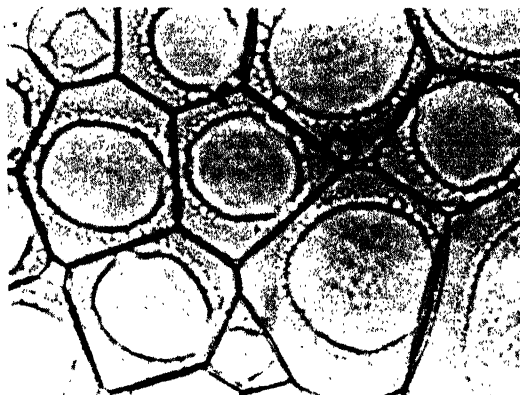
Fig. 41. Vacuolation in the complex coacervate adjacent to the wall (p. 469, Fig. 37b) after the action of 10 m. eq. per 1 $K_3Fe(CN)_6$ at constant pH.
a: a short time after the passage of the salt solution.
b: some time later
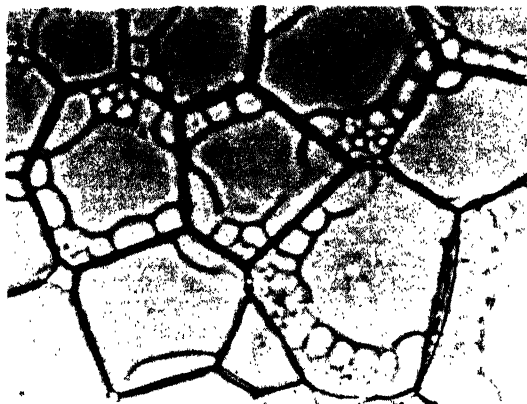
b

ation while at the same time the colloid content of the central vacuole must increase on account of increased solubility.

What happens after restoring the original pH (see Fig. 40) does in fact point to this The vacuolized state of the coacervate instead of decreasing immediately, persists for the time or even increases in intensity by the formation of a new generation of vacuoles. Furthermore a large number of small drops of coacervate separate out in the central vacuole on account of the decrease of the solubility of the coacervate.

Similar considerations hold in the explanation of the vacuolation phenomena when currents of salt solutions are made to flow past. Since salts make the coacervate richer in water (suppressing action p. 349, Ch. X, § 2f) one would not expect any .vacuolation. Nevertheless it occurs with various salts, as can be explained by the influence salts exert on the optimal mixing proportion of the colloids (p. 354-355 and p. 365). Indeed as regards the morphological pictures which occur (intensity.

of the vacuolation, localization of the vacuoles first appearing) the salts arrange themselves in the order of the continous valency rule: $3 - 1 .. 2 - 1 ... 1 - 1 ...$ $1 - 2 ... 1 - 3$ (p. 352-355, Ch. X, § 2f). The intensity of vacuolation increases in this series from left to right and the place in the coacervate where the vacuoles first appear shifts in the same direction from directly against the cell membranes towards directly against the boundary coacervate/vacuole. Compare for example the powerful action of $K_3Fe(CN)_6$ (salt type 1—3) in Fig. 41 (p. 473).

We cannot go into details here but we mention in addition the particular point that with the change of the salt concentration of the liquid flowing along the membrane a new generation of vacuoles frequently also appears in the coacervate, while at the same time the vacuoles already present become smaller or disappear.

c. Reversible colloid displacements between gelatinized parietal complex coacervate and central vacuole. Accumulation of neutral red[1].

If after the coacervation (pH 3.7) the complex coacervate (gelatin + gum arabic) has become parietal and one subsequently passes cold dilute buffer of pH 3.7 along the membrane, the parietal coacervate gelatinizes. We shall not go into detail here on the morphological changes which occur (for example gelation vacuoles and their properties).

If we add 1/200 % of neutral red to this buffer, then neither the gelatinized coacervate nor the vacuole takes up any microscopically visible colour (Fig. 42d).

If this buffer is replaced by a dilute buffer of pH 6 containing 1/200 % neutral red, then coloration of the vacuole gradually occurs. The latter first colours diffusely
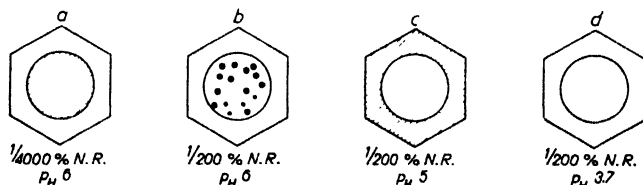
Fig. 42. Reversible colloid displacement between gelatinised complex coacervate on the wall and central vacuole (see text).
a: diffuse accumulation of neutral red in the vacuole.
b: "grain" accumulation of neutral red in the vacuole.
c: accumulation of neutral red in the gelatinised coacervate while the vacuole is uncoloured.
d: neutral red is not taken up either by the vacuole or by the gelatinised coacervate.

but presently dark coloured drops form in it (Fig. 42b). It now appears that it depends on the neutral red concentration what the final stage of the neutral red accumulation is.

With a very small concentration, for example 1/4000%, only diffuse accumulation occurs in the vacuole (Fig. 42a). The two states a and b of Fig. 42 can further be transformed reversibly into one another by changing the neutral red concentration in the liquid flowing along the membrane.

If we start from state b and now pass dilute buffer of pH 5 containing 1/200 % neutral red along the membrane, then the intensely coloured drops disappear, but the gelatinized parietal coacervate is coloured red (Fig. 42c).

[1] H. G. BUNGENBERG DE JONG and R. C. BAKHUIZEN VAN DEN BRINK, Proc. Koninkl. Nederland. Akad. Wetenschap., Amsterdam, 50 (1947) 436.

All four states *a*, *b*, *c* and *d*, can be transformed reversibly into each other by changing the neutral red concentration or the pH.

For an explanation it is in addition important that:
1. gum arabic coacervates with neutral red (type dicomplex coacervation colloid anion — micro cation);
2. gum arabic, if present alone in the cell compartments, accumulates neutral red at all the pH 's mentioned here, thus even at p·H 3.7;
3. the I.E.P. of the gelatin used amounts to 5.

The failure of any coloration to occur in *d* must be due to all the negative charges of the gum arabic being here completely taken up through complex relations by the positive charges of the gelatin. This is in agreement with the fact that the optimum mixing ratio of gelatin and gum arabic corresponding to pH 3·7 was enclosed in the cell compartments.

The gradual accumulation of neutral red in the vacuole at pH 6 indicates that the vacuole gradually begins to contain gum arabic as is understandable because at this pH the gelatin is already negative and thus can no longer conclude complex relations with the gum arabic (compare p. 465, § 4f).

Thus the gum arabic gradually diffuses into the vacuole, is even accumulated in it to a higher concentration than remains present in the gelatin layer (on account of the negative charge of the gelatin).

State *c* however does cause some surprise. Lowering the pH from 6 to 5 makes evidently all the gum arabic disappear from the vacuole. Thus a colloid displacement takes place to the gelatin gel layer although the latter is at or near its isoelectric point.

An explanation is given by the consideration that at the isoelectric point a protein is not really uncharged but has positively ionized groups next to negatively ionized ones. The macromolecules of the gum arabic can then attach themselves to these positive groups with some of their negatively ionized groups. The remaining negative groups of the gum arabic then remain free and can bind neutral red cations.

d. Changes in the wetting of the cell membranes

As already stated in § 5 a (p. 467) coacervates in general wet the celloidin membranes of the prismatic cells. This also holds for the coacervate which is produced after accumulation of a basic dye (for example toluidine blue) in the cell compartments filled with gum arabic sol. In the final state it wets the cell walls and lies spread out in a thin layer over them[1].

By passing a very dilute salt solution (e.g. NaCl 5 m. aeq. p. l) over (suppressing action on the dicomplex coacervate dye cation — arabinate) the dye is gradually washed out again and the ultimate equilibrium state will be that all coacervate has disappeared from the cell compartments and the latter are again filled with a gum arabic sol.

Now dynamic colloid morphology studies the path along which a colloid system proceeds from the one equilibrium state (here coacervate spread out over

[1] H. G. BUNGENBERG DE JONG and B. KOK, *Proc. Koninkl. Nederland. Akad. Wetenschap.*, Amsterdam, 43 (1940) 728.
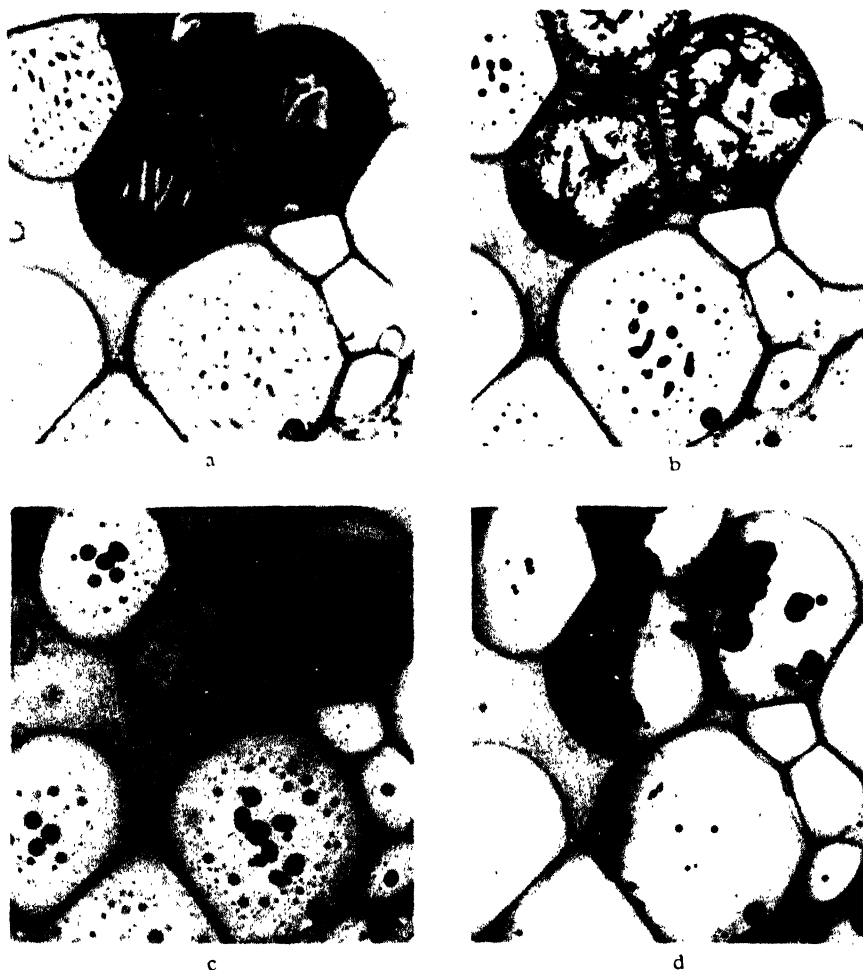
Fig. 43.   Changes in the wetting of the cell membrane (192 × lin.).
Many cells of the tissue have been injured whereby the gum arabic has wholly or largely left them.
Only three cells have not been injured and have accumulated toluidine blue to a maximum extent.

a: Initial state: the toluidine blue arabinate coacervate is on the wall. The white streaks and
    spots in these cells are places where because of folds in the celloidin membrane the coacervate
    on the wall is very thin and thereby transparent.

b and c: successive stages of the change in the wetting caused by treatment with 5 m. eq. per l
    NaCl. Pulsating vacuoles are formed in the coacervate drops in c.

d: In a later stage and after application of a direct current field (anode on the left, cathode on the
    right) see text.

the cell wall) to the new equilibrium state (here cell compartments filled with gum
arabic sol). Such a path does not in general need to be morphologically simple. In
our case this path is in fact not simple. One might, for example, suppose that the thin

coacervate lamella situated on the celloidin walls becomes gradually thinner and finally dissolves completely there.

In reality the path, along which the final state is reached, is much more complicated. Striking morphological pictures are produced which point to a decrease of the wettability of the walls by the coacervate.

The latter draws away from the cell walls and rounds itself off, after which the coacervate proceeds as separate spheres into the interior of the cell compartments (see Fig. 43 $a \rightarrow b \rightarrow c$).

We thus arrive at the conclusion that the coacervate surface already exhibits changed wetting properties under the prevailing circumstances — whereby the coacervate will ultimately dissolve completely. The coacervate drops, which have made their way into the interior of the cell compartments, further undergo the same striking changes internally, which one can study outside of the cell compartments on free coacervate drops (pulsating vacuoles p. 454, § 4a). Finally they dissolve completely and the cell compartment is still filled with a weakly coloured gum arabic solution. On further washing this colour also disappears and the final state is reached at which the cell compartment is only filled with a gum arabic sol.

With the dilute NaCl solution used the whole process of removal of toluidine blue proceeds slowly, because the relatively strongly bound dye cations must be replaced by the hardly bound Na-ions (constituting a diffuse double layer with the polyvalent arabinate ions).

As more and more dye cations are removed, the solubility of the coacervate will increase and its electropheretic charge will become more and more negative.

Thus already in Fig. 43c a certain amount of the arabinate toluidine blue coacervate is in solution, which follows from the darker background shade of the three cell compartments considered as compared to the same three in Fig. 43b.

That this diffuse coloration is due to dye cations still bound to the arabinate, appears from the fact that after application of a direct current field (fig. 43d) not only the remaining coacervate drops (apart from some still adhering to the cell wall) move rapidly towards the anode, but that also the blue general colloration in the cellcompartment shifts to the anode.

The phenomenon of pulsating vacuoles which accompanies the dissolving of the coacervate drops (Fig. 43c), seems also to be related to the negativation of the coacervate. Such a negativation also occurs under the circumstances described in § 4 a (p. 454, contact of the coacervate drops with excess of gum arabic sol).

In this connection we may remember that formation of foam bodies and hollow spheres (processes which seem closely related) is also induced by a sufficient negativation of the gelatin-gum arabic complex coacervate (see p. 459, § 4c and p. 463, § 4d).

§ 6.   COLLOID MORPHOLOGY OF PHOSPHATIDE COACERVATES

a.   Multiple coacervate drops in phosphatides

Phosphatide coacervates behave *discrepantly* in all kinds of respects[1] (see already p. 406, and for anomalous wetting behaviour p. 436). Some of the most striking anomalies will be discussed in this paragraph. With them among other things multiple coacervate drops occur, by which we understand colloid bodies which consist of several liquids

---

[1] H. G. BUNGENBERG DE JONG and R. F. WESTERKAMP, *Bioch. Z.* 248, (1932) 335.

placed inside each other. They were observed in the coacervation of phosphatide sols with $CaCl_2$ in the presence of organic liquids partially miscible or inmiscible with water. If for example the influence of increasing amounts of added aniline is investigated (Fig. 44) then with aniline concentrations, which still dissolve in the
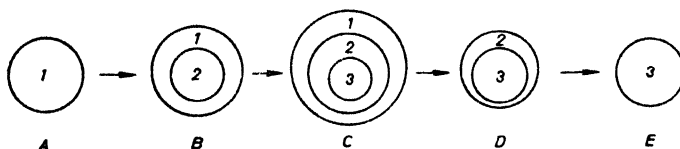
Fig. 44. Multiple coacervate drops (B, C, D) in phosphatide coacervates (see text).

water present, only flocculation occurs. With somewhat more aniline however thoroughly liquid homogeneous coacervate drops are produced (Fig. 44a). With gradually larger amounts of aniline, the multiple coacervate drops in question are now produced and indeed double or triple drops according to the amount added. (See Fig. 44b and c).
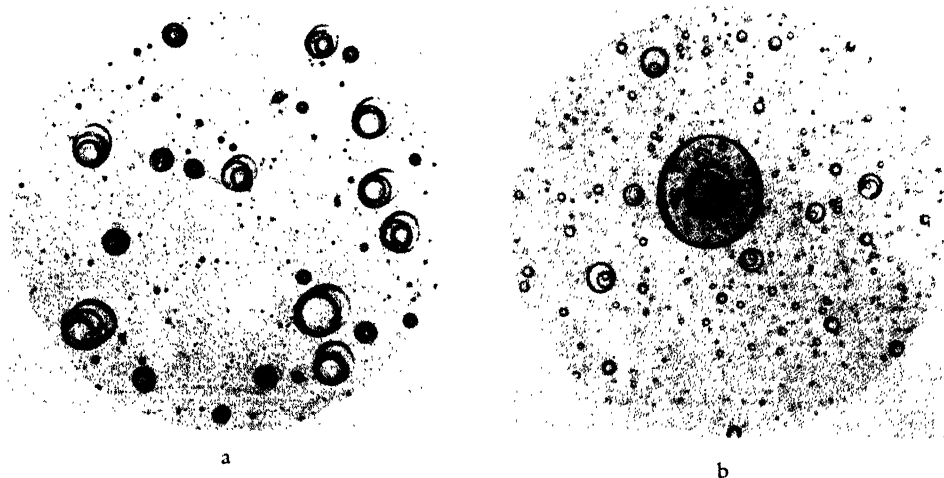
Fig. 45. Examples of triple coacervate drops in phosphatides (78 λ lin.).
a: in a soya bean phosphatide sol containing $CaCl_2$ to which 0.4 cc of aniline were added per 10 cc.
b: in an egg lecithin sol containing $CaCl_2$ in which 30% triolein and 30% oleic acid calculated on the egg lecithin are present. The triple coacervate drops are thoroughly liquid because 3.6 mol per l ethyl alcohol was present in the medium.

It is probable that shell 2 is also a coacervate [1] which can therefore coexist with the peripheral coacervate. On the other hand it has probably no meaning to regard liquid 3 as a coacervate but as an aniline-rich liquid which contains dissolved phosphatide. With a sufficient excess of aniline only 3 remains (see Fig. 44e), while

---

[1] Thus in it vacuolation and streaming occur in a D.C. field as in shell (1) although of much smaller intensity.

in between there probably exists the state in which double drops, composed of 2 and 3 are formed (Fig. 44D).

We are inclined to consider the two coexisting coacervates as belonging to one of the two media partially miscible with one another. Thus coacervate 1 as belonging to the medium consisting mainly of water, coacervate 2 as belonging to the medium consisting mainly of aniline. These coacervates then so arrange themselves morphologically that they are added as concentric shells between the surrounding water-rich medium and the centrally situated aniline-rich medium.

The relatively large solubility of coacervate 2 in the aniline medium then causes the displacements in the morphological picture when the amount of aniline added is slightly increased.

Thus the phosphatide is successively displaced from the sphere of the aqueous medium (coacervate 1 disappears) completely towards that of the aniline medium (coacervate 2 also disappears).

The occurrence of multiple coacervate drops is not restricted to the example given. They can also occur under certain circumstances in the presence of substances insoluble in water such as oleic acid, trioleine, etc. See Fig. 45, which gives an example of this sort for triple coacervate drops.

b. Behaviour of phosphatide coacervate drops in a d.c. field

Phosphatide coacervates[1] also behave in an anomalous manner with regard to their behaviour in an electric field[2] (see Fig. 46). They do indeed exhibit similar vacuolation and streaming phenomena to those described on p. 348 and 452, but the phenomena are here modified by the particular nature of the coacervate surface
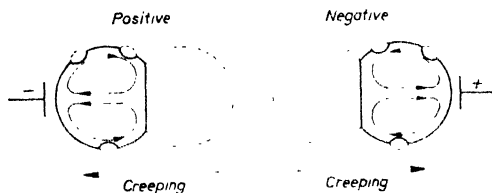
Fig. 46. Behaviour of phosphatide coacervate drops in an electric field as a function of the sign of the electrophoretic charge (see text).

(see p. 480, § 6c). Small vacuoles frequently cannot make their exit but remain as protruding vesicles stuck in the surface of the drop. Since these mark the surface it can be observed with certainty from their movement that the surface of the coacervate drop itself moves in the electric field in the same sense as the coacervate content lying close under the surface. These attached vacuoles are transported to one side of the drop

---

[1] The ordinary coacervates are meant, that is to say those which we have indicated in § 6a as „belonging to the aqueous medium", thus for example A in Fig. 44, to which in the mean time coacervates also belong which are obtained with the aid of all kinds of non-electrolytes miscible with water + salts (p. 405).

[2] H. G. BUNGENBERG DE JONG and R. F. WESTERKAMP, *Bioch. Z.* 248 (1932) 335.

and there coalesce with other vacuoles already transported there to form a vesicle continually increasing in size.

The vesicle wall is however very thin and not or hardly visible but its presence is betrayed by the flattening on one side of the coacervate drop. These coacervate drops can while lying on a surface also carry out creep movements in an electric field in the direction of the arrow below the figures. Compare with Fig. 20 (p. 452).

c. The anomalous behaviour of the vacuoles in phosphatide coacervates

Thoroughly liquid coacervate drops[1] of phosphatides (lipophilic) have the property in common with those of extremely hydrophilic colloids that they readily coalesce on contact with each other. With regard to vacuoles however an anomalous behaviour occurs in the phosphatide coacervates (in typical cases).

This behaviour is set out in the following survey and, where useful, the behaviour of the extremely hydrophilic, non-lipophilic coacervates is added between brackets.

a. If two vacuoles in the coacervate drop come into intimate contact with one another then a "facetted" doublet (Fig. 47a) is produced with an invisible or very difficulty visible partition at the position of the articulation [normal: the two vacuoles coalesce to a new spherical vacuole].
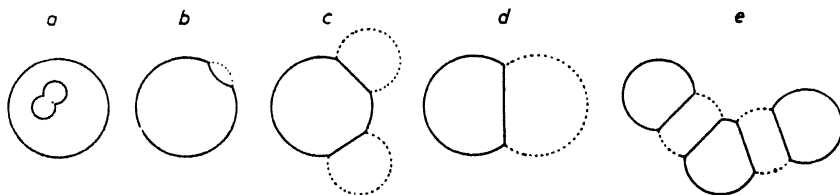
Fig. 47. Peculiar behaviour of the vacuoles of phosphatide coacervates (see text)

b. If a vacuole comes into contact with the inside of the surface of the coacervate drop then an indentation is formed there while the membrane, of the vacuole now imprisoned in the surface of the drop, which protrudes into the surrounding medium is again frequently not visible (Fig. 47b). Its position can however be readily established by adding a little indian ink to the medium [normal: the vacuole breaks through, ejects its contents into the surrounding medium, after which the surface of the coacervate drop rounds itself off again].

c. Vacuoles which have been fixed in this way in the outer surface of the drop, can indeed coalesce together sideways into a larger vacuole fixed in the same way.

---

[1] Here also the „ordinary" coacervates are meant, see note on previous page.

If this takes place to a considerable extent then the coacervate drops assume the shape of spheres from which segments have been cut off normally here and there. (Fig. 47c). The attached vacuoles also behave in the same way in an electric field (p. 479, Fig. 46) but since they can be transported only to one side of the drop by the flow of the coacervate surface and then coalesce sideways, the coacervate drops assume the shape of a halfmoon (Fig. 47d). In these cases also the membrane extended into the surrounding liquid from the now very large vacuole attached sideways is frequently invisible.

d. If two or more coacervate drops with their vacuoles attached sideways come into contact with each other, then the vacuoles frequently run together so that the coacervate drops have now been joined together by invisible or difficultly visible membranes. It is characteristic of such systems, which are displaced as a whole by convection currents, that two contiguous drops present parallel faces towards each other (Fig. 47e).

e. On destruction of the above described invisible or difficultly visible membranes (for example by moving a fine glass needle to and fro at some distance to the right of the flat face of the drop in Fig. 47d by means of a micromanipulator) the contents of the attached hyaline bubble are found to be completely miscible with the surrounding medium. Immediately after the rounding off of the coacervate drop — evidence of a successful destruction — one does not see anything happen at the place where the vacuole was situated. Vacuole contents and surrounding medium are therefore completely miscible.

We can thus summarize the above in two points:

1st   very thin phosphatide films exist which can separate two aqueous, completely miscible, or even identical liquids (the content of two vacuoles, see point a) or almost identical liquids (vacuole contents and surrounding medium),

2nd   their production is initiated by coacervation and subsequent vacuolation.

For an explanation of the nature of these peculiar films it is of importance to start from the contrast between ready fusibility of coacervate drops with each other and the formation of stable films on the contact of vacuoles with each other or of vacuole with the inside of the drop surface. It already follows from this that the boundary between coacervate and aqueous medium (surrounding liquid or vacuole liquid) has different properties on the two sides. If indeed two such boundaries come into contact with one another with their surfaces directed towards the aqueous medium, coalescence takes place (coalescene of two massive coacervate drops or the case mentioned in point *d* of the above survey).

On the other hand a stable film is formed when these boundaries encounter one another with their surfaces directed towards the coacervate side (*a, b* and *c*). If one assumes that there is a film of orientated phosphatide molecules at the coacervate/ aqueous medium boundary, then it is understandable without further argument that the two sides of this film bear very different characters. This makes it probable that the very peculiar films described above are bimolecular.

The question still remains open whether in this bimolecular film the two mono-

molecular sheets lie with their hydrocarbon ends opposite each other or with their polar groups[1].

For a further study of the abnormal phenomena described in this § it seems of interest that they all (see p. 478, Fig. 44 and 45; p. 479, Fig. 46 right; p. 480, Fig. 47) have been recently found to occur also with oleate coacervates under specified conditions[2].

[1] To the question of how the phosphatide molecules are orientated in the coacervate/aqueous medium boundary one is inclined to answer that the polar portion of the molecule will indeed be directed towards the aqueous medium, and consequently the non-polar portion towards the coacervate side. In that case the bimolecular film is set on the two outer sides with polar groups.
There are however indications which make one doubt whether the assumed orientation is really correct. If one assumes the opposite orientation then the polar portions would be adjacent to one another in the bimolecular film whereby the electrical dipoles can form precisely a favourable salt pattern. For a detailed elaboration of this idea see:
H. G. BUNGENBERG DE JONG and J. BONNER, *Protoplasma*, 24 (1935) 198 and a shorter summary in *Proc. Koninkl. Nederl. Akad. Wetenschap.*, Amsterdam, 38 (1935) 797.
In these publications it was still only supposed that the double film consists exclusively of phosphatide molecules (+ possibly cholesterol molecules). As well as these phosphatidic acids are also present in the phosphatide coacervates which makes them just sensitive to cations (see p. 406).
Double films derived from the above, in which these elements are also included, are appropriate as models of the protoplasmic membrane.
See H. G. BUNGENBERG DE JONG and G. G. P. SAUBERT, *Proc. Koninkl. Nederland. Akad. Wetenschap.*, Amsterdam, 40 (1937) 295 and *Protoplasma*, 28 (1937) 352.
[2] To be published in *Proc. Koninkl. Nederland. Akad. Wetenschap.*, Amsterdam.

XII. GELS

P. H. HERMANS

Laboratory for Cellulose Research of the AKU and Affiliated Companies, Utrecht

---

§ 1. INTRODUCTION. DEFINITION OF THE GEL STATE

At the beginning of a treatise on gels it seems necessary to consider the question as to what the concept "gel" shall embrace.

Twenty years ago, D. JORDAN LLOYD[1] began her survey on the problem of gel structure with the words: "The colloidal condition, the gel, is one which is easier to recognize than to define". In earlier days colloid science has often built its terminology on the macroscopic aspects of matter long before a physico-chemical understanding of structures or even of the phenomena involved was reached. Therefrom has arisen that systems of similar appearance or apparently analogous phenomena have been designated by the same nomenclature. Moreover, the evolution principle has often been applied. Let us quote some examples related to the subject "gels".

The system formed if a fairly concentrated sodium silicate solution is neutralised and kept for some time without stirring, a coherent transparent microscopically homogeneous mass, is a typical gel. Upon drying this gel, a hard transparent glasslike substance is obtained, which was still termed a gel. The hard mineral agate, likely originating in nature from a silicic acid gel, has, consequently, been considered as a gel. Is it surprising that THOMAS GRAHAM, the creator of the term gel, was inclined to consider ordinary glass as a gel too?

Typical gels are formed if substances like rubber, gelatin, polystyrene and cellulose are allowed to swell through absorbing fluids; the original dry substances have therefore also been classified as gels.

If a solution of sodium silicate is neutralised and vigourously stirred, a voluminous flocculent precipitate is obtained. It has been termed a "gelatinous" precipitate and was usually considered as being a gel. So have been other precipitates of similar appearance e.g., the hydroxides of the heavy metals. Amongst precipitates there are, however, many gradations in appearance and consistency. Then doubt may arise as to whether the term gel applies. The precipitates of the silverhalides, are they gel-like? The flocculae formed upon coagulation of a colloidal gold solution, are they gels?

Difficulties of classification like these frequently occur in classical colloid chemical terminology and they become acute anew as soon as a better general insight into

---

[1] D. JORDAN LLOYD, The problem of gel structure, in J. ALEXANDER's *Colloid Chemistry*, Vol. I, New York 1926, p. 767—782.

the physico-chemical background of the phenomena is reached. If the latter allows of a more precise definition, one should not hesitate to endeavour a revision and to break with tradition where this seems to be desirable. It would seem that we shall have to face such a situation in the beginning of this chapter with regard to the concept gel and we shall base ourselves on the principles set fourth in the chapter General Introduction (Vol. I, Chapter I). According to the latter, we shall assign the term gel to systems characterized by the following points:

(a) they are coherent colloid disperse systems of at least two components;

(b) they exhibit mechanical properties characteristic of the solid state;

(c) both the dispersed component and the dispersion medium extent themselves continuously throughout the whole system (cf. also p. 2).

In contrast to (a) and (b), (c) is not a criterion based on phenomenology but refers to the structure which cannot be recognized at first sight. Nevertheless, it seems desirable to add this criterion to the definition of gels, since a number of systems conforming with (a) and (b) which should not be considered as gels however, can, by direct observation (e.g., with the ultramicroscope) be identified as not conforming with (c).

Solid colloidal systems like ruby glass (which may be considered as a solidified colloidal suspension of individual gold particles in molten glass) or like a "solid foam" are excluded by this definition.

It is true that recent investigations have shown that possibly there may exist systems of a gel-like character not corresponding to (c). The coherence of the dispersed component is here due to long range forces. Such systems seem to occur in the tobacco mosaic virus and perhaps also in certain soap jellies. We shall however abstract from these particular cases.

As a criterion for "properties characteristic of the solid state" we may conveniently ask the question whether or not the system exhibits a *yield value* upon mechanical deformation, e.g., if exposed to a shearing tension, and, hence, whether it shows the phenomenon of strain.

There is no doubt that the definition of the concept gel offered here applies to typical gels like gelatin, agar, cellulose and silicic acid gels which are also characterized by the fact that one of the components of the system is a low molecular fluid. Like almost any definition (particularly in colloid science!) difficulties may arise when considering certain limiting and transitionary cases. Some of these will be referred to below.

The gel state is very common in the domain of macromolecular substances and may, in a sense, be considered as being connected *inseparably* with their solid state[1].

The condition gel is, however, by no means confined to macromolecular systems. A great many gels have been prepared also from other substances and perhaps we may even say that almost any substance can be transferred into a gel-like condition with the aid of suitable operations.

Though this volume treats the macromolecular aspects of colloid science and though we shall, hence, mainly focus attention on macromolecular gels in this chapter, it seemed necessary to include also a discussion of gels in general, since this subject has not yet been treated elsewhere in this book, and, for more than one reason a separate treatment seemed to be undesirable.

---

[1] Apart from few exceptions which will be mentioned later.

Before concluding this section the term *Xerogel* introduced by FREUNDLICH may be recalled. Typical macromolecular gels consisting of a macromolecular component and a fluid shrink if the liquid component is evaporated or otherwise disappears. They then transform into a dense solid mass entirely consisting of the macromolecular component. This macromolecular solid of amorphous or cryptocrystalline nature may be called a resin or, in comparison to other amorphous solids, a macromolecular glass. It does not fit the definition of a gel as formulated above, but is so intimately connected to the gel state (since it is formed from a gel and can be easily transformed into a gel by adding a swelling agent), that we may consider it as a boundary condition of the gel state in macromolecular systems. Therefore, the term xerogel[1] seems to be justified. Examples are dry gelatin, dry cellophane, polystyrene and rubber. Natural products with a particular biostructure like native cellulose fibres, starch and leather, showing the phenomenon of swelling, may be also considered as xerogels. On the other hand, substances like well cured bakelite exhibiting no sorptive or swelling power to any appreciable extent shall preferably be classified as resins. The term xerogel should be confined to sorbing and swelling systems[2].

§ 2   SOME BRIEF HISTORICAL NOTES IN CONNEXION WITH TERMINOLOGY[3].

Earlier theories of gel structure were manifestly influenced by the particular horizon of the investigators and the particular objects investigated.

VON NäGELI[4], focussing attention on natural objects like cell walls, starch grains and the like, postulated a discontinuous granular structure of very small crystalline particles, carrying round themselves concentric shells of tightly bound water. The birefringence of these objects would be due to orientation of these crystallites termed *micelles*. In colloidal solutions of cellulose and starch these particles would be dispersed through the liquid phase and gelatination would be due to the formation of a frame work structure throughout the liquid consisting of numerous mutually cohering micellae, binding on their surface a portion of the liquid.

The micellar theory in its essential features was supported by numerous authors at the time of the vivid revival of colloid science in 1910 and after, e. g., by ZSIGMONDY (1911), BACHMANN (1912), ARISZ (1915), BRADFORD (1918) and others. This theory may be classified as the solid-liquid theory of gel structure.

Wo. OSTWALD (1909) put forward a liquid-liquid theory which has but historical interest.

The point of view generally accepted at the present moment is a similar form of the solid-liquid theory, postulating that both the solid and the liquid components are continuous in themselves. P. P. VON WEIMARN (about 1910) held similar ideas, entirely fitting in with modern views. One of the earlier forms of this theory may be called the fibrillar theory. The solid component is assumed to be of a more or

---

[1] From ξηρός = dry.

[2] Though it seems to have been shown that in molten glass upon cooling transitionary structures occur which fit the definition of a gel (see page 655), the solidified glass should, hence, not be termed a xerogel.

[3] Cf. D. JORDAN LLOYD, loc. cit.

[4] VON NäGELI, *Pflanzenphysiologische Untersuchungen*, Zürich 1858; Cf. OSTWALDS, *Klassiker der exakten Wissenschaften*. No. 227; A. FREY, *Die Micellartheorie*, Leipzig 1928.

less fibrillar nature and to form a continuous framework or meshwork throughout
the system. This idea has been especially suggested by the study of gels formed by
certain low molecular substances which, under other conditions, can also crystallise
from their solutions (see p. 491). The fibrillar theory was first suggested by VAN
BEMMELEN (1898). In BÜTSCHLI's honeycomb idea ("Wabenstruktur") a continuous
structure of the solid component and a discontinous dispersion of the fluid were
assumed, the latter being enclosed in polygonal cels formed by the former.

The solid-liquid theory has given rise to much discussion on whether the solid
component is crystalline as NÄGELI had postulated, and, if not, which is its condition.
We shall see that the problem is more or less irrelevant with respect to the general
conception of the condition, the gel, since not all gels have the same architecture.
Either alternative and a continuous series of transitory states may be realized. It is
to be noted that there is less essential difference between VON NÄGELI's ideas and
the modern concepts of gel structure than the brief characterization given here might
suggest. As a matter of fact, upon reading V. NÄGELI's more than half a century old
papers, one is impressed by the far reaching and advanced ideas contained therein.

A number of investigators have put forward the solid solution theory. PROCTER,
investigating the swelling of gelatin gels, was, as early as 1914, the first who decided
in favour of this theory and postulated a frame work of molecular dimensions, certainly
an admirable achievement in those days. The solid solution theory was also supported
by KATZ in 1918.

Valuable contributions to the clarification of gel structure and the terminology
concerned were recently made by FREY-WYSSLING. He proposed to conserve the
classical NÄGELI-terminology in a modified form, designating the typical gel structure,
consisting of two dispersed components each continuous in themselves, as *micellar
systems*, in contrast to the common *disperse systems* of classical colloid science, consisting
of individual dispersed particles in a *dispersion medium*. A scheme, as given in one
of his recently appeared books[1], illustrating the general terminology of both doctrines
and the transitionary states between the two is reproduced here.

	Dispersion doctrine	Transitionary states	Micellar doctrine
colloidal conditions	sols	"gel solutions"	gels
colloidally dispersed substance	individual particles, acting as kinetic units (possibly: macro-molecules)	mutual interaction of the particles; association of macromolecules	coherent frame work (micellar structure)
solvent	dispersion medium		imbibition fluid
structure	not structured	beginning structure	structured
elasticity	not elastic	structural viscosity	elastic; yield value
condition	fluid	viscous fluid	solid

[1] A. FREY-WYSSLING, *Submikroskopische Struktur des Protoplasmas und seiner Derivate*, Berlin
1938, p. 86; 2nd Edition in English, Elsevier, New-York, Amsterdam 1948, p. 52.

This scheme very clearly points out the essential features of the subject. The nomen-clature proposed by FREY-WYSSLING, which is very attractive indeed, implies, however, that the stem word "micelle" should be dropped in all other constructions. This consequence will be hardly welcomed by all investigators and there will certainly be many objections.

True, the different meanings attached by various authors to the term micelle have already given rise to a great deal of confusion in some branches of colloid science. The word micelle has e. g., been used for:

1. any polymolecular dispersed colloid particles,

2. polymolecular dispersed colloidal particles of a crystalline nature exclusively,

3. complex colloid dispersed particles consisting of the dispersed component including shells of solvent adhered to it or including solvent of imbibition,

4. any dispersed particle of colloidal dimensions, also if it consists of only one macromolecule,

5. crystalline particles of colloidal dimensions, no matter if dispersed in a solution, or, if occurring in a coherent solid system,

6. crystalline regions in xerogels and gels of macromolecular systems not representing individual particles, but merely regions of higher order in an otherwise amorphous aggregation of macromolecules.

In connexion to the study of swelling phenomena and chemical reactions in cryptocrystalline macromolecular systems, the terms *intramicellar* and *intermicellar* processes were introduced by KATZ and have become very popular. The former refers to processes occurring between the crystallites, the latter to occurrences inside them. They can be discriminated by X-ray examination. (See section 6b. 8, p. 577).

In the scheme proposed by FREY-WYSSLING intramicellar processes are those occurring inside the solid framework structure and intermicellar ones those occurring in its interstices. It would seem, however, that there are good arguments to maintain these concepts in the sense originally proposed by KATZ, since they are then subject to an unequivocal criterion which may be read from the X-ray diagram.

This example and others, not to be discussed here, show anew that it is very difficult, if not impossible, to find a system of terminology which is entirely consistent in itself.

A general agreement between scientific writers, though very desirable, has not yet been reached and various authors continue to use the word micelle in various meanings. In this chapter we shall, however, avoid the use of this term, except in the derivations *inter*micellar and *intra*micellar, in the sense, as originally introduced by KATZ.

Various suggestions have been put forward as to the subdivision of gels according to certain properties.

HARDY (1899) has discriminated heat-reversible and non heat-reversible gels according as the gel can be liquified and solidified at will by a change in temperature. Many examples of heat reversible gels occur, such as gelatin in water, agar-agar in water, methyl-cellulose in water. The gels of cellulose in water are, however, not heat-reversible and there is no valid argument to classify them in another group. This subdivision has, therefore, little sense.

FREUNDLICH (1922) classified gels as elastic and rigid ones. The former can be

reswollen after having been dried, and the latter are then devoid of swelling power. He therefore preferred later to discriminate swelling and non swelling gels for the same groupings. The swelling gels form a more compact mass upon drying than the other group which transform into a more or less porous solid. Perhaps we can then better speak of *reversible* and *irreversible* gels, taking as a criterion whether the gel, after having been freed from the volatile component by evaporation, reswells or not upon readdition of the solvent. In practice this classification leads to almost the same grouping as the subdivision: *macromolecular* and *non-macromolecular gels*, which, at the present day, seems to be the most rational one.

Finally, mention may be made of the classification *heterogels* and *isogels*, introduced by Wo. OSTWALD. The latter refers to the particular case that the solid component is the polymer of the monomeric liquid component (Cf. Chapter II by R. HOUWINK p. 40).

We shall, however, not lay much stress on general classification, since the latter is no essential condition for dealing with the fundamental aspects of our subject.

§ 3. THE FORMATION AND STRUCTURE OF GELS

3 a. General condition of gel formation

Gels may be formed either from a solution or from a solid substance (a xerogel!) exhibiting swelling power. The latter instance is confined to macromolecular substances. No such restriction exists for the formation of gels from solutions. At the beginning of the experiment the solution may be either a colloidal solution or an ordinary solution of a low molecular substance. We shall treat separately the formation of gels from solutions and the formation of gels by swelling of a macromolecular substance. In this section, the former subject will be comprehensively dealt with, the latter merely briefly in a general sense, since a more detailed treatment of the swelling proces will be given in § 6 on Sorption and Swelling.

Two fundamental conditions must be fulfilled in order that a gel be formed from a solution.

1. a solid substance shall separate from the solution in a finely dispersed „colloidal" state,

2. the separated solid particles shall neither be deposited by gravity nor remain in a colloidal suspension as freely moving kinetic units, but they shall join together to form a continuous coherent framework throughout the mass of the solution[1].

From any process, virtually capable of producing the separation of some solid phase, either crystalline or amorphous, from a solution (or a melt) gel formation may result. Hence a condition of super-saturation caused either by a change of temperature, by evaporation, addition of another substance (a non solvent or a salt) or by a chemical reaction, occurs as the first step. Secondly, suitable conditions allowing of the formation of a continuous pattern or frame work of colloidal fineness must be realized[2].

---

[1] The French use the characteristic expression: „prise en masse" in this connection.

[2] This is the very idea already held by v. NÄGELI, who wrote for instance: „Wir können uns dieses Gelatinieren wohl nur in der Art vorstellen, dass die Mizelle sich in Ketten aneinander einhangen und ein Gerüst von Balken mit weiten Maschen bilden *(Theorie der Gärung*, 1879, p. 102, 127).

A schematic picture of four different possibilities concerning the nature of the framework, which was already shown in Chapter I of Volume I, is again reproduced in Fig. 1.

Fig. 1a shows spherical particles adhering to each other in more or less linear arrangement, Fig. 1b rodlet shaped particles building up a similar continuous framework. In Fig. 1c the case of linear macro-molecules forming a framework consisting of molecular chains with junction points of a crystalline nature is represented. Fig. 1d shows a typical case of gel formation by chemical cross-linking of dissolved linear macromolecules. Such a case can, e. g., be realized if a rubber solution is vulcanised
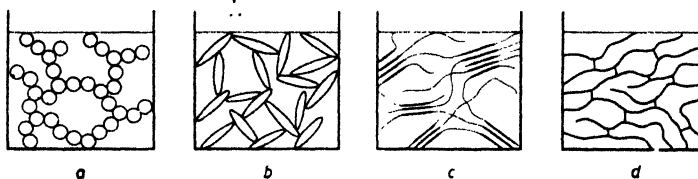
Fig. 1.   Four types of gel structure (diagrammatic).

in situ. This can e. g., be done by simultaneously introducing hydrogen sulphide and sulphur dioxide gas into the solution. The junction points then formed between the molecular chains make, in a sense, the whole quantity of rubber present in the solution to one colossal macromolecule remaining dissolved as before, but with loss of the fluid character of the solution, since the different parts of the solution are now devoid of the capacity to flow freely along each other. A similar structure emerges from the polymerisation of monomeric styrene in solution upon addition of a small amount of divinylbenzene (cf. p. 165).

An interesting boundary case is the polymerisation of pure liquid styrene. The process then continuously passes through the "isogel" state (a frame work of poly-styrene molecules soaked with the monomer) to the xerogel state, which is reached it the entire amount of styrene is polymerised.

The places where the particles, originally moving as kinetic units in the solution, adhere to each other will be called the *junction points* of the gel. The forces causing interlinking may be of various nature, comprising ordinary cohesive forces (VAN DER WAALS forces), polar forces, heteropolar and homopolar valency bonds. Besides direct material contact between the original particles, other molecules or atoms may parti-cipate in the formation of the links. In a number of cases, bridges formed by tightly 'bound water molecules may e. g., play a part; cohesion is then of a polar nature.

In the discussion of gel structure MANEGOLD[1] has considered the question of the volume fraction of the space occupied by a continuous frame work. If the particles are spherical, a regular packing can be reached in various ways, varying the coordination number of each particle from 3 to 12.

The volume fraction occupied by the spheres in the densest packing (coordi-nation 12) is 0.74 and in the least dense packing (coordination 3, see Fig. 2) is 0.056. Hence, in the latter case, only 5.6% of the volume is occupied by the particles. This is even considerably less than the volume occupied by the molecules of a gas at its critical temperature, which is 0.083. In many gels, however, the relative volume

[1] E. MANEGOLD, *Kolloid-Z.*, 96 (1941) 186.

of the framework is still very much smaller. For the vitreous body of the eye, it is about 0.001. It would seem therefore that a system of cohering more or less spherical particles comes into consideration only for gel-like systems with a relative low content of the liquid component, like gels of bentonite and $BaSO_4$.

From these arguments it follows that the ordinary gels must contain fibrillar systems consisting of strongly anisodiametric structural units. As MANEGOLD showed, a lower limit for the relative volume of a frame work structure of this kind does not exist.
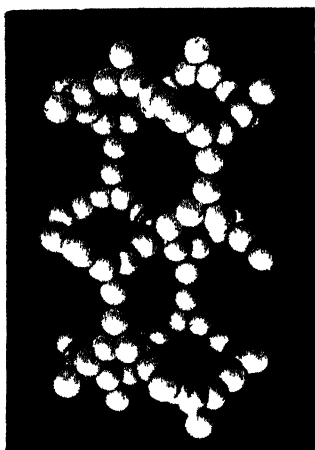
Fig. 2.   Packing of spheres after HEESCH and LAVES[1]

For any system capable of gel formation there will be, however, a limit of dilution where gel formation becomes impossible. If then conditions occur bringing about demixture, flocculation is observed instead of gel formation. A coherent frame-work can no longer be formed. There is, however, no sharp limit between gel formation and the formation of a flocculatous precipitate. The occurrence of a "prise en masse" or the separation of a precipitate largely depends upon the conditions of the experiment. At high dilutions stirring or shaking of the solution will have a marked bearing on the appearance of the system after demixture (see also § 5a, p. 509).

In many cases it is even very difficult to discriminate between a gel and a liquid. Very diluted gelatin solutions do not gel upon cooling, as more concentrated solutions do. Flocculae do not appear either in this case. In other words, the dilute solutions maintain an apparent fluid character, they can be poured from one vessel into the other. With very sensitive means, howeve., it can be shown that, nevertheless, such solutions exhibit a yield value, i.e.. they do not flow below a certain shearing force. (cf. p. 505). In a sense, such flowing systems can be also considered as gels, since they have a property in common with solids, i. e., the property of exhibiting a yield value. The junction points of these gels can, however, easily be broken. Upon standing the gel-character reestablishes itself. Solutions of this kind may be designated as thixotropic, though this term is mostly used in connexion with the reversible liquefaction of more solid gels which can not be poured from one vessel into the other.

It might be stated that almost all gelling systems become thixotropic below a certain dilution (see also § 4f and § 5b).

b. Gelformation and demixture

It will be clear that gel formation from a solution is always connected with some process of demixture. A new phase of solid nature must tend to separate from the solution, but this new phase must remain divided in particles of colloidal dimensions and the particles must, in some way, adhere to each other in order to form a continuous network throughout the liquid.

A particularly important form of demixture leading to gel formation is crystal-

---

[1] Z. Krist., (A) 85 (1933) 443

lisation. A solid phase separates from the solution and the crystals form a continuous frame work. This may occur in true solutions of either molecular or macromolecular substances. (Cf. section c and d).

Gels may also emerge from colloidal solutions containing solid particles, either of a crystalline or of a non crystalline nature. If the solution is subjected to some treatment, capable of bringing about "coagulation", the solid phase may either appear as a flocculous precipitate or, under suitable conditions, also as a gel of one of the types diagrammatically represented in Fig. 1a and Fig. 1b. This is why the addition of salts to concentrated sols of the type of the ferric hydroxide sol may cause gel formation and why in such cases the conditions of gel formation may follow the SCHULZE-HARDY rules (cf. Volume I, General Introduction).

Even the separation of a fluid phase may give rise to gel formation. A well investigated example of this kind were the gelatin gels studied by HARDY [1], with the purpose of elucidating gel structure. He prepared solutions of gelatin in a mixture of equal volumes of water and ethylalcohol. They were perfectly transparent above 20°. On cooling below this temperature the solutions became cloudy and microscopical examination revealed that very small droplets separated from the solution. These droplets consisted of what we call at present a *coacervate* (cf. Chapter VIII). Upon further cooling, the droplets solidified and adhered together, thus building up an anastomising frame work of spherical particles hanging together in linear, often branching, rows. The whole liquid then solidified.

Upon cooling more rapidly, the droplets became smaller and smaller and finally microscopically unresolvable structures were obtained. If the concentration of the gelatin exceeded 36 %, not the droplets separating solidified, but the liquid in which they were dispersed. The gel formed consisted of a solid structure formed by the latter and in which numerous liquid droplets remained enclosed. These experiments, however, only show that a gel of gelatin *may* be formed in this way and have the structure set forth here. They do not prove that all gelatin gels belong to this type. The former represent typical systems of composite nature, in the sense as explained in Chapter I of this volume.

The formation of gels of the type of Fig. 1a has also been microscopically observed by HARDY (*loc. cit*). In coagulation of egg albumin he observed the separation of spherical particles of 0.75-1 μ diameter, adhering together to form an open network with polygonal meshes.

Gel-like systems can also emerge from concentrated suspensions of coarser particles, if these, for some reason or another, adhere together in such a manner that much liquid is enclosed. If the cohesion between particles building up the framework of the gel is only weak and can be broken by mechanical disturbance, re-establishing itself when the mechanical action ceases, we have the phenomenon of *thixotropy* (cf. § 5b, p. 510).

c. Gel formation from crystallisable low molecular substances

Every chemist knows the phenomenon of abundant crystallisation from a solution whereby a coherent mass of (usually anisodiametric) crystals fills the bulk of the solution, the mother liquor being wholly enclosed by the crystal cake. This may be regarded as a macroscopical image of gel formation from a crystalline substance. If the dimensions of the crystals (in one or two dimensions at least) remain in the colloidal region, a gel would be the result. Factors favouring the formation of a) a great number of very small crystals instead of a few large ones, b) an irregular anisodiametric shape of the crystallites, will tend to favour gel formation. In order that a great number of small crystals be formed, there must be an abundant formation of crystallisation germs.

Extensive and elucidating researches on gel formation from a number of inorganic

---

[1] W. B. HARDY, *J. Physiol.*, 24 (1899) 288; *Proc. Roy. Soc. London*, 66 (1900) 95.

substances we owe to VON WEIMARN[1]. At the time his work seems to have been less noticed and appreciated than it deserved; it contained the very basis of the modern concepts of gel structure and gel formation.

VON WEIMARN experimented mainly on the formation of precipitates of slightly insoluble inorganic compounds like silver sulphate, barium sulphate, aluminium hydroxide, which he prepared by double decomposition of other salts, added together under varying conditions of concentration and temperature and with or without addition of other liquids, like alkohol, to influence solubility and medium conditions. The essential result of his experiments may be cast into the form of the following equation[2]

$$n = kP/L \qquad\qquad (1)$$

where L is the solubility of the solid which precipitates,

P the degree of super-saturation (measured by the difference between the concentration present at a given moment and the normal saturation concentration,

k a factor involving the viscosity,

n the number of crystallisation centers (germs) generated.

P and k are functions of temperature.

Crystallisation is largely governed by two factors, the number of crystallisation centers formed and the velocity of crystallisation.

Generally speaking, the number of germs and the crystallisation velocity will be small in very dilute solutions and low degrees of supersaturation. Few, slowly growing, crystals will be formed, but in length of time the crystals may become larger and larger if a small degree of supersaturation is maintained. In concentrated solutions the velocity of growth increases and, if a substance of low solubility L is formed by a reaction in concentrated solutions of highly soluble reactants, the degree of super-saturation P will be large; n will, hence, be at maximum and an enormous number of very small crystals will be formed.

Further it is a well known fact that such conditions also favour the formation of irregular ramified crystal forms like dendrites, trichospherites and the like. According to MANEGOLD[3] twin formation should be hold as a frequently occurring cause of branching and extraordinary crystalline forms. If those growing crystalline particles are very near to each other, they will easily give rise to a cohering framework throughout the liquid. This, doubtlessly, is the nature of the gels prepared from crystallisable substances. Mostly, these gels are not very stable. A more or less rapid recrystallisation may set in and the gel is transformed into a mass of microscopically visible crystals. Upon evaporation of the solvent from gels of this kind a solid mass is obtained, which is not capable of swelling upon readdition of the solvent.

Examples of such gels are those obtained by mixing 5 n solutions of barium

[1] P. P. VON WEIMARN, *Kolloid-Z.*, 1 (1907) 76; 2 (1908) 199, 230, 275, 301, 326; 5 (1909) 122; 6 (1910), 182; 9 (1911), 25; *J. Russ. Chem. Soc.*, 40 (1908) 1787; 46 (1914) 110, 624; 47 (1915) 2163. Cf. the survey written by this author in J. ALEXANDERS, *Colloid chemistry*, New York 1926, p. 27—101; *Kolloides und Kristalloides Lösen und Niederschlagen*, Dresden 1925.

[2] VON WEIMARN's reasoning is only very briefly and compendiously sketched here. It is worth while to read his own survey (*loc. cit*).

[3] E. MANEGOLD, *Kolloid-Z.*, 96 (1941) 186.

sulphocyanate and manganous sulphate or by allowing a strong solution of sodium alcoholate in alcohol to rapidly absorb gaseous hydrochloric acid. Gels of barium sulphate and sodium chloride are then obtained.

In the course of time, a great many crystalline organic substances have been discovered which, upon recrystallisation from a suitable solvent, form gels instead of depositing crystals. Examples are dibenz-oyl-l-cystine [1] (upon pouring its alcoholic solution in water), camphoryl phenyl thio semicarbazide [2] (by recrystallisation from toluene); Azomethine [3] (upon recrystallisation from organic solvents), 1. cyclohexyl cyclohexane cis 1.2 diol [4] (upon recrystallisation from water), Lithium urate [5] and others.

All these gels are unstable and transform themselves more or less rapidly into
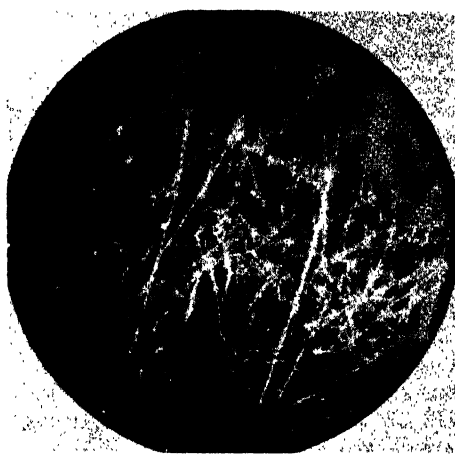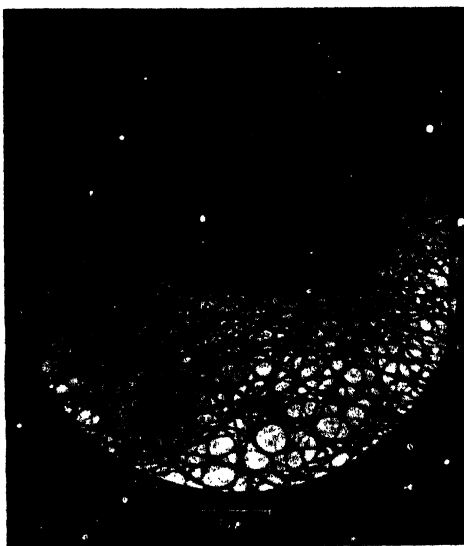
Fig. 3a.   Dark field micrograph of an optically resolvable patch in the gel of 1 cyclohexyl cyclohexane cis 1.2 diol. Magnification 260 ×   .

a solid mass consisting of separate crystals. Ultramicroscopical examination of these gels has shown that they exhibit a fine fibrillar structure. Often transitions between ultravisible and crystalline structure have been observed. Next to transparent patches of the gel, such with an optically resolvable fibrillar structure may occur. A micro photograph of a resolvable patch of the gel of 1 cyclohexyl cyclohexane cis 1.2 diol is reproduced in Fig. 3a. It can be seen from the photograph that the substance tends to crystallise in strongly anisodiametric, almost threadlike forms. The thickness of these threads vary from microscopical dimensions to very thin no longer resolvable ones.

Fig. 3b.   Electron-microscopic image of a V_2O_5 gel. 20.000 × photographic magnification included 50.000 ×.

[1] R. A. GORTNER and W. F. HOFMANN, J. Am. Chem. Soc., 43 (1921) 2199.
[2] E. HATSCHEK, Kolloid-Z., 11 (1912) 158; Trans. Farad. Soc., 12 (1916) 17.
[3] W. B. HARDY, Proc. Roy. Soc. London, 87 (1912) 29.
[4] P. H. HERMANS, Kolloid-Z., 97 (1941) 231.
[5] H. SCHADE and E. BODEN, Z. physiol. Chem., 86 (1913) 238.

In Fig. 3b, a microphotograph of a vanadium pentoxide gel, taken with the electron microscope, recently published by FREY-WYSSLING and MÜHLETHALER[1], is reproduced, showing a coherent recticular structure which must be visualized as having a spatial extension. (The focus depth of the electronic microscope being very much larger than that of the ordinary microscope, the image represents a relatively thick layer of the structure projected in one plane). As compared to the length of 1μ indicated the larger pores are of the order of magnitude of 150 m μ.

It will be clear that in the formation of very small crystallites of irregular shape less latent heat of crystallisation will be liberated than in the formation of larger well formed crystals. HARDY found that in azomethine gels a certain quantity of heat has to be withdrawn from the solution in order to bring about crystallisation, but that a further quantity has to be withdrawn before visible crystals separate. He further found that azomethine solutions may be super-cooled with respect to gelation and that such super-cooled solutions may be caused to gel by slight warming or sowing with already formed gel. The presence of nuclei (crystallisation germs) seemed to be essential for gel-formation. At re-crystallisation crystals separated outside the gel in the solution at the cost of the gel, melting away at its surface.

In a number of cases, concerning freshly prepared gels from inorganic substances, no crystalline structure could be detected by X-ray examination. Under varied conditions (as precipitation at higher temperatures or ageing) they may or may not yield a crystalline diffraction pattern. The former case is met with in the trivalent iron and aluminium hydroxides, the latter in those of quadrivalent thorium, zirconium and cerium. The precipitates of the hydroxides of bivalent zinc and magnesium always exhibit a crystalline spectrum. These facts first discussed by HABER[2] and later further investigated by BÖHM et al[3] are not contradictory to the crystallisation theory, set forth hereabove. Crystallites may be so small as to escape detection by X-ray examination and there will also be continuous transition between well defined crystallites, those with more and more distorted lattice formation, and other aggregations, resulting from the very same forces of intermolecular attraction, which ought to be considered as being amorphous. Upon formation of the latter, less latent heat will be evolved. Just the same they represent the products of an intrinsically identical process of demixture[4].

d. Gel formation from solutions of macromolecular substances

According to modern views, the gels formed from solutions of linear macromolecules may in many instances also be considered as due to a process of crystallisation or at least to something very similar to crystallisation. For obvious reasons a recrystallisation to form larger microscopically visible individual crystals does, however, not occur in macromolecular gels.

The conditions governing gel formation are again those which reduce the solubility as e. g., change of temperature, addition of a non solvent to the solution, a

[1] A. FREY-WYSSLING and K. MÜHLETHALER, Vierteljahresschr. Naturf. Ges. Zürich, 89 (1944) 214.

[2] F. HABER, Ber. 55 (1922) 1717.

[3] J. BÖHM and W. P. NIKLASSEN, Z. anorg. Chem., 132 (1924) 1; 149, (1925) 214.

[4] Cf. HABER's explanation as to why the readiness of crystallite formation decreases according as the valency of the metal ion becomes greater (loc. cit.).

chemical reaction modifying the character of the macromolecules. If the solubility limit is surpassed, the macromolecules tend to cohere to each other. Since the macro-molecules are very long, one molecular chain may form cohesive junction points with other molecules at several places simultaneously and it is clear that this will give rise to the formation of a molecular network structure. (Fig. 4 en 5)[1].
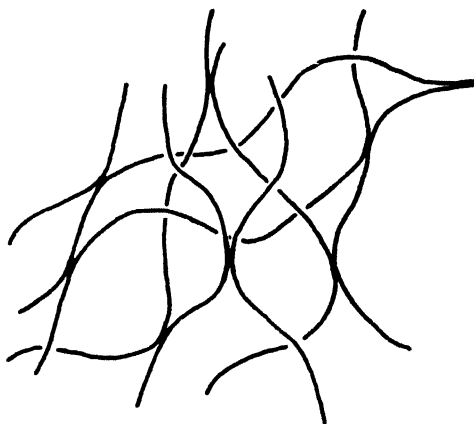
Fig. 4.   Molecular network.
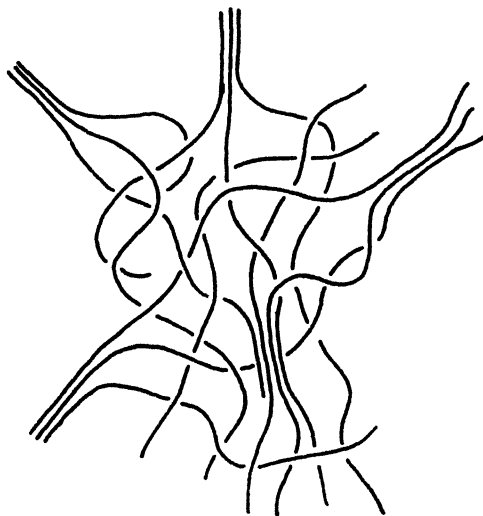
Fig. 5.   Molecular network with crystalline junction points.

If the junction points are small and comprise only a small number of chains, the gel formed remains a typical homogeneous one phase system. Most macromolecules, however, tend to crystallise. Then, some junction point will consist of several molecules arranging themselves in lattice order (see Fig. 5). Just as in low molecular systems, crystallisation wil be governed by the degree of supersaturation and the number of crystallisation centres formed. Several observations, which will be discussed later, (Section 4a) show that actually these factors also play a part in gelation. If crystallisation centres are formed, the crystalline regions will tend to grow, but, as will be easily understood from Fig. 5, the mutual entanglement of the chains will soon put an end to the process. This is why macromolecular gels always remain cryptocrystalline. Thermodynamically, a large monocrystal of the macromolecular substance would be the most stable endpoint. This endpoint will, however, never be reached. True, gels of this kind show the phenomenon of ageing. In the course of time, changes occur in the gel which lead to assume that recrystallisation occurs to a certain extent and larger crystallites are formed, but they always remain below the border line of resolvability in the microscope.

---

[1] The possibility of a molecular network has already been put forward by Wo. OSTWALD *Kolloid-Z.*, **40** (1926) 58; **69** (1934) 339; **80** (1937) 375 and W. W. HALLER, *ibid*, **56** (1931) 257.

If the crystallites surpass certain dimensions, the system of Fig. 5 might be considered as a two phase system. It is more or less arbitrary at which limit of crystallite dimensions the system is no longer monophasic. Neither it can be said which is the lower limit of the dimensions of a crystallite and what is the size of a crystallisation germ. Perhaps it is practical to say that we will only speak of crystallites if the x-ray diffraction pattern actually shows distinct interference spots due to a crystalline lattice. The number of chains in a crystallite which corresponds to this criterion is very much greater than in our diagrammatic picture of Fig. 5 and will amount to, say, 30 as a minimum. This figure is a rather striking one if the chains in the original solution were randomly distributed and assumed to have randomly kinked forms (cf. Chapter IV). Yet it would seem that in many cases the numer of chains united in a crystallite is often still con-

Fig. 6.   Chain molecules with low distance order.

siderably larger than 30. The formation of crystallites of that size is not easy to imagine. It seems to imply either a rather high mobility of the chains or that the actual arrangement of the chains in solutions of moderate and high concentration is not a quite random one in very small regions. It has more than once been suggested that we have to deal with a "low distance order" in such solutions.

Even in ordinary low molecular liquids neighbouring molecules are not orientated wholly at random, but are more or less ordered locally in much the same way as in a crystalline lattice. This order extents over small regions only, and orientation continuously varies in direction when we move from one point inside the liquid to the next. Low distance order is enhanced if the molecules have the shape of rodlets, which is demonstrated by the conditions giving rise to the formation of fluid crystalline phases in organic substances

Fig. 7.   Diagrammatic representation of the structure of a gelatin gel according to GERNGROSS and HERRMANN.

and also by the model experiments of STUART[1]. Macromolecular solutions, before gelation, may exhibit small distance order[2] (see diagrammatic representation

[1] W. KAST and H. A. STUART, Z. angew. Chem., 53 (1940) 53; H. A. STUART, Naturwiss., 31 (1943) 123. (Also see p. 120).
[2] Cf. O. KRATKY, Kolloid-Z., 68 (1934) 347; 96 (1941) 301; P. H. HERMANS, Kolloid-Z., 83 (1938) 71.

in Fig. 6). Then the formation of larger crystallites is more readily conceivable.

The first to propose a gel structure of the kind discussed here were GERNGROSS and HERRMANN [1]. A diagrammatic picture of the structure of a gelatin gel given by these authors is reproduced in Fig. 7. Similar ideas have been put forward by various other authors with regard to cellulose gels and other objects [2]. Further support to the theory of crystallisation is afforded by the observation that gelation is facilitated by inoculation of the solution with particles of the gel [3].

It will be clear that the dimensions of the crystallites in gels of this kind will vary according as the conditions of gelation and the age of the gel change. According to similar rules as those governing the size and growth of crystals in general, rapid gelatination will yield another distribution of the crystallite size than a slow one. The age of the gel, the temperature at which it is kept, the addition of foreign substances and other factors will have a bearing on its architecture and especially on the number and dimensions of the crystalline junction points. Maintaining our above mentioned definition of a crystallite, we shall also have to take into account the existence of
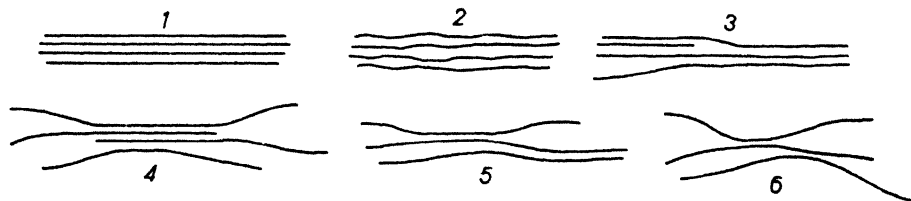
Fig. 8.   Various types of junction points arranged in the order of decreasing cohesive energy (junction point spectrum).

a series of junction points which can no longer be termed crystallites, but who are yet not less important with regard to structure and properties of the gel. Junction points down to a simple permanent contact between the monomeric groups of adjacent chains may be thought of.

In view of later considerations, it is useful to introduce the concept of the *junction point spectrum* of the gel, comprising the entire series of junction points present, arranged according to increasing cohesive energy. It may be thought to be represented by a curve, giving the fractional numbers of junction points as a function either of their cohesive energy, or of their size.

In Fig. 8 a diagrammatic representation of various types of junction points (numbered in the order of decreasing cohesive energy) is given.

[1] O. GERNGROSS, and K. HERRMANN, *Biochem. Z.*, 228 (1930), 409; *Z. physik. Chem.*, 310 (1930), 371; *Kolloid-Z.*, 60 (1932) 276.

[2] W. T. ASTBURY, *Trans. Faraday Soc.*, 29 (1933), 193; F. D. MILES, *ibid.*, 29 (1933) 110; F. T. PEIRCE, *ibid.*, 29 (1933) 50; A. FREY-WYSSLING, *Protoplasma*, 25 (1935) 261; 26 (1936) 45; O. KRATKY, *Kolloid.-Z.*, 70 (1935) 14; 84 (1938) 149; E. SAUTER, *Z. physik. Chem.*, B 36 (1937) 129; P. H. HERMANS and A. J. DE LEEUW, *Kolloid-Z.*, 81 (1937) 300; *ibid.*, 83 (1938) 71; E. BAUMGARTNER, *Thesis*, Bern, 1944.

[3] S. N. BANJERI and S. GOSH, *Z. anorg. Chem.*, 1947 (1930) 130.

There will always be a gradual transition from the simplest chain-to-chain-contacts to crystallites of increasing size and lattice order. Only if the latter surpass certain dimensions and if the condition of 3-dimensional lattice order is sufficiently fulfilled, will they be detectable by X-ray examination. All these junction points will, however, have a bearing on gel formation and on the mechanical properties of the gel.

Some changes to which the gel is subjected, owing to varying reactions or conditions like swelling, contraction, temperature change and so on, may be expressed in terms of a shift of its junction point spectrum.

Macromolecular gels formed from a solution by cooling will again liquefy upon raising the temperature (e.g., gelatin and agar-agar gels in water). Sometimes the gel is formed by heating the solution and becomes liquid on cooling (methylcellulose in water, cellulose nitrate in alcohol). This may be connected respectively with a positive and a negative temperature coefficient of solubility. The temperature of melting is different from that of solidification. We shall revert to these phenomena later (§ 4a).

If gelation is due to a chemical reaction as e. g., the formation of a cellulose gel by decomposition of a cellulose xanthate solution with an acid, we may speak of an irreversible sol-gel transformation.

Gels formed by the addition of a non-solvent (in aqueous solutions this may also be an electrolyte) will redissolve in the original solvent.

Besides gel formation due to crystallisation or junction point formation, another type of gel formation may occur in macromolecular systems, whereby cross-linking of the molecules by primary valence bonds is the essential process.

A typical example is the polymerisation of styrene to which small quantities of divinylbenzene have been added (cf. p. 165) the latter giving rise to a certain degree of threedimensional polymerisation and to the formation of a network structure[1]. Recently, SIGNER and TAVEL[2] have described gels made from solutions of cellulose and cellulose derivatives by the action of substances like the chlorides of dicarbonic acids whereby chemical cross-links are evidently formed.

Gels of this kind are obviously examples of monophasic gels. They are always characterized by limited swelling power in all solvents, since the primary valence cross linkages cannot be opened by solvent action alone. Such gels also always show a "memory" for the volume which they originally had at their formation (cf. p. 527).

A particular example of a macromolecular gel is the silicic acid gel, which is formed from sodium silicate solutions upon acidification. According to WILLSTäTTER et al.[3] mono or disilicic acids are formed in the reaction, which undergo polymerisation giving rise to high molecular silicic acids[4]. Since the polymerisation is a 3-dimensional one, a molecular framework is built. Polymerisation continues long after the solution has set to a gel. The case is comparable to that of the polymerisation of phenolformaldehyde solutions setting to a gel (cf. Ch. II, p. 39).

---

[1] H. STAUDINGER and W. HEUER, Ber., 67 (1934) 1164.

[2] R. SIGNER and P. V. TAVEL, Helv. chim. acta, 26 (1943) 1972.

[3] R. WILLSTäTTER, H. KRAUT, and L. LOBINGER, Ber., 61 (1928) 2280; 62, (1929) 2027.

[4] Cf. also C. B. HURD and P. C. MERZ, J. Am. Chem. Soc. 68 (1946) 61, where the earlier literature is cited.

§ 4. GENERAL PHYSICAL PHENOMENA OF THE SOL-GEL TRANSFORMATION

a. Transition temperature

It has already been stated that, in thermoreversible gels, the sol-gel transition temperature is no precisely defined constant and that it depends on the rate of cooling, and further that hysteresis phenomena occur. An 0.5% agar-agar solution sets to a gel when cooled to 35°. When heated, it liquefies again, but not until a temperature of 90° is reached. Below 90° the gel does not liquefy, however long it may be maintained at the particular temperature. Consequently, it is possible to obtain this system either in a sol or a gel condition at any temperature between 35° and 90°. This is an extreme case, but most of the thermoreversible sol-gel transformations show hysteresis to such a degree, that the sol or the gel state can be obtained within a temperature of 10 to 20 degrees[1].

This phenomena might be explained thus. A certain degree of undercooling is necessary to bring about solidification and crystallisation. The size of the crystals increases after the solidification has begun. The melting point of the crystals largely depends on their size. One may also say that the "melting point" of the junction points depends on their extension and their place in the junction point spectrum[2]. Upon reheating, the junction points are loosened according to their place in the spectrum. The largest ones may resist temperatures above the original solidification temperature of the gel.

ARISZ[3] and DE BOER and DIPPEL[4] have found that the temperature of liquefaction lies the higher the longer the gel has been aged after its solidification. DERKSEN[5] showed that certain changes in the X-ray diffraction patterns of gelatin gels, indicating a disordering of the crystalline lattice, occur at a higher temperatures in aged than in freshly prepared gels[6].

A very striking example has been reported by ARISZ[7]. A 10% gelatin solution in a glycerol-water mixture (with 32% glycerol) was cooled from 70° tot 44°. Kept at this temperature, the solution set to a gel. If the same solution was cooled to 35° and allowed to solidify, a gel was obtained which, upon heating at 44°, first liquefied but later set once more to a gel. The experiment succeeded only if the 35° gel was fresh. Gels kept during longer time at 35° did *not* liquefy at 44°. This may be interpreted thus: rapid undercooling to 35° gives rise to a great number of small crystals melting below 44°. The temperature of 44°, however, still lies below the temperature of demixture of the solution. After some time, crystals can be formed at this temperature too. They are larger than the ones originally formed at 35°. If the 35°-gel

---

[1] E. HEYMANN, *The sol-gel transformation*, Paris 1936.

[2] The term "melting point" has here a similar meaning as, e. g., the melting point of glaubersalt crystals under water.

[3] L. ARISZ, *Kolloidchem. Beih.*, 77 (1915) 12.

[4] J. H. DE BOER and C. J. DIPPEL, *Rec. trav. chim.*, 52 (1933) 216.

[5] J. C. DERKSEN, *Thesis*, Amsterdam 1934.

[6] It is also known that raw rubber, which freezes, i. e., crystallises, upon cooling shows a higher melting point according as the crystallisation was allowed to proceed further; A. VAN ROSSUM and J. LOTICHIUS, *Z. Kautschuk*, 5 (1929) 2.

[7] L. ARISZ, *loc. cit.*

is kept longer at 35° and crystallisation has had time to proceed and to produce larger crystallites; melting below 44° does not occur.

ARISZ also found that ageing takes place more rapidly at higher than at lower temperatures. Gels kept at 0° aged slower than those kept at 35°. This is quite analogous to the behaviour of undercooled ordinary solutions of crystallisable substances. DERKSEN, who made extensive X-ray studies on the crystallisation of gelatin gels, found that recrystallisation in gels kept at very low temperatures proceeded slower than in gels kept at room temperature. These examples may suffice to support our explanation of the phenomenon of hysteresis.

b. Thermal effects

In several cases it could be shown that heat is given out upon gel formation from a solution. This heat may be interpreted as the latent heat of crystallisation. The evolution of heat is a gradual one, just as in cooling of a saturated solution of an ordinary substance with a positive temperature coefficient of solubility. In macromolecular solutions, however, the cooling curves do not show sudden breaks, since the process of crystallisation is, so to say, smeared out over a broader range, because a spectrum of crystallite sizes, having continuously varying heats of crystallisation, is involved.

A distinct difference between the cooling curve of gelatin solutions and that of pure water was established by LOTTERMOSER and MATTHAES[1]. A positive heat effect was also found by other authors[2]. HARDY showed that heat is evolved upon gel formation from azomethine solutions[3]. Similar results have been obtained for gelling soap solutions[4].

In concentrated gels showing thixotropy, thermal effects have not been observed. Such gels are not macromolecular systems. Gelation is, here, not connected with crystallisation but involves much weaker forces. Absence of a thermal effect is, hence, not surprising.

c. Optical effects [5]

In many cases the setting of a sol to a gel is accompanied by an increase of the TYNDALL effect (light scattering), indicating an increase of particle size [6]. This is what should be expected in macromolecular gels, if crystallites or junction points of larger size than the lateral dimensions of the molecules are formed. From observations on the intensity and the depolarisation of the TYNDALL light in gelatin solutions (cf. Vol I, Ch. III), DONNAN and KRISNAMURTI (loc. cit.) concluded that the original

---

[1] A. LOTTERMOSER, and M. MATTHAES, Z. physik. Chem., 141 (1929) 129.

[2] K. WINKELBLECH, Z. angew. Chem., 19 (1906) 1260; W. B. PLEASS, Proc. Roy. Soc. London, A 126 (1930) 406; L. J. W. HOLLEMAN, Thesis, Leiden 1932, showed that heat is absorbed when gelatin gels are liquefied.

[3] W. B. HARDY, Proc. Roy. Soc. London, 87 (1912) 29.

[4] M. H. FISCHER, Kolloid-Z., 46 (1929) 359.

[5] For a more detailed survey cf. E. HEYMANN, The Sol-Gel Transformation, Paris 1936.

[6] Gelatin: E. HATSCHEK and R. H. HUMPHREY, Trans. Faraday Soc., 20 (1924) 18; L. ARISZ, Kolloidchem. Beih., 7 (1915) 1; F. G. DONNAN and K. KRISNAMURTI, Colloid Symp. Monogr., (1930) 1; E. O. KRAEMER and S. T. DEXTER, J. Phys. Chem., 31 (1927) 764; Cellulose esters: F. W. MARDLESS, Trans. Faraday Soc., 18 (1923) 1.

anisometric particles in the sol formed aggregates more spherical in shape. This can be well understood on the basis of the picture of crystallite formation. KRAEMER and DEXTER (*loc. cit.*) found that gelatin solutions of hydrogen-ion concentrations larger and smaller than the isoelectric point, show no change in the TYNDALL-effect during the gel formation. On the other hand, even in very diluted isoelectric gelatin solutions an increase of the TYNDALL-effect on cooling is observed, although no solidification occurs in this case. The dependence of the TYNDALL effect on the pH may be understood from observations of DERKSEN[1], who found by X-ray examination that the quantity of water absorbed in the crystallites (this water can be compared with water of crystallisation) increases according as the pH of the solution differs from the isoelectric point. The difference between the optical density of the crystallites and that of the surrounding solution, which determines the magnitude of light scattering, will consequently be maximum at the isoelectric point, where the water content of the junction points and crystallites is a minimum. An optical phenomenon which has drawn much attention and is designated as the "mutarotation" of gelatine may be briefly mentioned here. The rotatory power of gelatin solutions towards polarised light is practically constant at temperatures between 80° and 35° but increases considerably at lower temperatures. As we shall see later, gelatin solutions at low temperatures are no longer true solutions but assume properties of gels; it would seem that the phenomenon of mutarotation is intimately connected with that of gel formation. Changes of pH and addition of salts influence mutarotation in the same sense as they do gel formation. The change in rotatory power may perhaps be explained by changes in the steric configuration of the chains which take place when the chains begin to cohere and form larger multimolecular aggregates [1a].

d. X-ray examination [2]

The existence of crystallites in gels of macromolecular substances was revealed by X-ray diffraction photographs in various instances. Well known examples are gels of cellulose and several of its derivatives, starch, silicic acid and gelatin. We shall not go into the details of the X-ray investigations, since we shall have to revert to this subject later (Sections 6b 8, p. 577 and 8c, p. 602).

In contrast to the opinions formerly held and dating back as far as C. VON NÄGELI, the crystallites occurring in natural cellulose fibres are completely disintegrated if the cellulose is dissolved [3]. Cellulose gels prepared from such solutions always yield X-ray diffraction patterns revealing the presence of crystallites, which, hence, must be formed by recrystallisation. In cellulose gels the amount of crystalline substance seems to be always approximatively the same and to depend little, if at all, on the conditions of preparation. It also remains sensibly constant upon drying of the gel [4]. This peculiar fact remains to be explained.

---

[1] *loc. cit.*, p. 499 and 502.
[1a] Cf. E. O. KRAEMER and J. R. FANSELOW, *J. Phys. Chem.*, 29 (1925), 1169; 32 (1928), 894.
[2] The literature on this subject is very extensive. We may refer to the book of J. R. KATZ, *Röntgenspektrographie als Untersuchungsmethode*, Berlin — Vienna 1934. The principles of X-ray examination in colloidal systems have been set forth in Volume I, General Introduction.
[3] O. KRATKY, *Kolloid-Z.*, 96 (1941) 301; G. CENTOLA, *Atti X. Congr. Int. chim. Roma*, 4 (1938) 117, 129, 138, 722, 728; Cf. also H. L. BREDEE, *Kolloid-Z.*, 94 (1941) 81; further M. MATHIEU, *La gélatination des nitrocelluloses*, Paris 1936.
[4] O. KRATKY and A. SEKORA, *Kolloid-Z.*, 108 (1944) 169; cf. P. H. HERMANS, *Contribution to the physics of cellulose fibres*, Amsterdam - New York 1946; *Bull. Soc. Chim. Belg.* 57 (1948) 123.

Silicic acid gels whose formation is not due to a process of crystallisation (cf. p. 498) but to a chemical cross linking, yield an amorphous X-ray diffraction pattern[1]. A faint indication of crystallisation is only found in aged dehydrated gels, particularly after the latter have been exposed to higher temperatures.

The crystallinity of gels prepared from acetone solutions of cellulose nitrate depends on the concentration of the solution from which they have been prepared and on the temperature at which they are kept[2].

Particularly well investigated is the crystallisation of gelatin[3]. Air dry gelatin samples yield an X-ray diffraction pattern exhibiting crystalline interferences next to broad diffuse rings, revealing that a part of the substance is in the amorphous condition. A concentrated gelatin solution in water shows the X-ray spectrum of an amorphous substance or a liquid. If cooled until gel formation occurs, it develops an X-ray spectrum with distinct crystalline interferences[4]. Upon liquefaction of the gel by heating it to 55°, these interferences disappear, to reappear upon cooling and gelation. The phenomenon was closer investigated by DERKSEN[5]. He endeavoured to measure systems in equilibrium and followed what happened, if a gel containing 42% gelatin was slowly warmed from 18° to 46°. Before the gel liquefied, the X-ray pattern showed characteristic changes; the crystalline interferences gradually became less sharp and intense. One of the characteristic interferences changed more rapidly than the other[6]. The last traces of the crystalline character of the gel disappeared at a temperature slightly over that at which liquefaction of the gel was complete[7].

It seems that the liquefaction of the gel coincides with the "melting" of the crystallites, though a portion of the latter, perhaps those crystallites which had the largest size in the original gel, may for some time still exist in the liquid state. It is not surprising that the gel may be liquefied before all the crystallites are molten. The coherence of the frame work may be destroyed after the greater part of the junction point spectrum has been cancelled. The gradual changes observed by DERKSEN may be interpreted as continuous shift of the latter. This is in conformity with the fact that also other properties of the gel change according as the temperature is raised. DERKSEN showed that the modulus of elasticity of the gel gradually decreased[8].

The liquefaction temperature of gelatin gels depends upon the concentration of

---

[1] S. KYROPOULOS, Z. anorg. Chem., 99 (1917) 197.

[2] M. MATHIEU, La gélatination des nitrocelluloses, Paris 1936, p. 44 Compare also § 8c, 1ð, p. 611.

[3] S. C. BRADFORD, in J. ALEXANDER's Colloid Chemistry, New York 1926, p. 761 mentions that small crystalline spherites of gelatine in sizes up to 3 μ may be obtained under conditions corresponding to a minimum degree of supersaturation and nucleus formation. (Very slow evaporation of very dilute aqueous solutions). It seems to be not very well established, however, whether the experiments eventually refer to a very low molecular fraction of the product.

[4] W. ABITZ, O. GERNGROSS, and K. HERRMANN, Z. physik. Chem. A, 150 (1930) 257; 165 (1931) 161; J. R. KATZ, J. C. DERKSEN, and F. W. BON, Rec. trav. chim., 50 (1931) 725, 1138; 51 (1932) 513, 835.

[5] J. C. DERKSEN, Thesis, Amsterdam 1935. Cf. J. R. KATZ, J. C. DERKSEN, and W. F. BON, Rec. Trav. chim., 50 (1931) 725, 1138; J. R. KATZ and J. C. DERKSEN, Collegium (1932) 931; J. C. DERKSEN, ibid. (1932) 838.

[6] Compare the changes in the X-ray diagram of cellulose nitrate upon swelling mentioned on page 549.

[7] Cf. also O. GERNGROSS, K. HERRMANN, and E. LINDEMANN, Kolloid-Z., 60 (1932) 281.

[8] Cf. also J. L. OUWELTJES, Thesis, Amsterdam 1942.

the solution from which the gel was prepared. According to SHEPPARD and SWEET [1] the liquefaction temperature falls from about 42° for a 48% gel to 26° for a 1% gel. According to X-ray observations, more and larger crystallites are present in concentrated gels than in diluted ones. The different junction point spectrum makes it clear that the temperature range of the sol-gel transformation will be different.

Finally it should be noted that the process of liquefaction can be also termed a process of dissolution. In fact, both concepts are equivalent in this instance, just as in the case of glauber salt crystals "melting" in water at the transition temperature of 32°.

The reverse process, the setting of a gelatin solution to a gel was also carefully examined by DERKSEN. If a 42% solution, showing an amorphous spectrum at 60°, was cooled at 22° and set to a gel, the crystalline spectrum reappeared and, kept at this temperature, its intensity and sharpness gradually increased for some days. He showed by dilatometric measurements that, simultaneously, a volume contraction occurred which ceased when the intensity of the X-ray spectrum reached its maximum. If the gel was rapidly heated to 35° (while it still retained the solid state), kept at this temperature for some time and then was rapidly cooled to 22°, the volume continued to decrease again for some time. This shows that reversible changes of the junction point spectrum of the gel and of its degree of crystallinity occur upon temperature changes.

The facts dealt with in this section may serve to illustrate that many even very complicated phenomena observed in the sol-gel transformation as well as the behaviour of gels in the realm of macromolecular substances, which for many years have puzzled the investigators, can be satisfactorily interpretated and often correctly predicted (qualitatively at least) with the aid of the picture of gel structure developed in § 3d) [2].

e. Volume changes

We shall forbear from treating this subject in detail and confine ourselves to the statement that volume changes, either a contraction (gelatin + water) or a dilatation (methyl cellulose + water), have been frequently observed to accompany the sol-gel transformation [3]. Since this is a common feature of any process of crystallisation or dissolution, no particular importance is to be attached to these phenomena. In the gelatination of thixotropic 9% sols of iron hydroxide, where gel formation is not due to a process of crystallisation, no volume change was observed.

In silicic acid gel formation, a considerable increase in volume is observed, starting directly after the acidification of the sodium silicate solution and before the solution has set to a gel. It also continues for a long time after setting. The volume increase here obviously runs parallel with the polymerisation of the silicic acid, a reaction involving the formation of water which was formerly chemically bound by the silicic acid. A dilatation is always observed in such condensation reactions.

[1] S. E. SHEPPARD and S. SWEET, Ind. Eng. Chem., 13 (1921) 423.

[2] DERKSEN was not yet aware of this explanation. The first who clearly put forward ideas of this kind was J. L. OUWELTJES, (loc. cit., p. 507). See also E. BAUMGARTNER, Thesis, Bern 1944.

[3] Cf. HEYMANN, loc. cit., p. 499.

f. Flow properties

It is clear that the viscosity of a gelling solution increases and reaches a very high value after the gel has set. If the initial solution shows anomalous (e.i., non NEWTONian viscosity, fluidity dependent on shearing force) the degree of departure from NEWTONian flow increases too. If the solution showed NEWTONian flow, its viscosity will become anomalous. Moreover a yield value appears, the liquid becomes elastic[1].

A great deal of work has been devoted to these phenomena[2]. It would seem, however, that this is not a place to enter into the details of the subject. We shall confine ourselves to the general features.

There seems to be little doubt that departure from NEWTONian flow is always due to some form of particle aggregation. In true solutions of macromolecules this aggregation may, for instance, be connected with association, if the average molecular weight in the liquid state or in solution exceeds that of the individual chemical molecule, e.g., if double or triple molecules occur. In such cases an equilibrium subject to the ordinary laws of mass action exists between the single molecules and the associates. The lifetime of the associates is finite and there is a permanent formation and decomposition of the complex particles.

Similar associations will frequently occur in solutions of linear macromolecules, especially in concentrated solutions, and if the chains bear polar groups. These associations will have the same character as those occurring in low molecular substances, the difference being, however, that one molecular chain, in virtue of its large extension, may eventually take part in association complexes with neighbouring ones at several places simultaneously. In other words, junction points of limited life-time will be formed throughout the liquid, giving rise to the formation of a kind of continuous network structure like that occurring in a gel. Owing to the ephemeral existence of the junction points, the system, however, does not assume the properties of a solid. It will still yield to even very small shearing forces, though with a speed depending on the average life-time of the junction points. Physically speaking, its viscosity will be proportional to the product of the shearing tension and the relaxation time of the junction points[3].

Any factor favouring association will enhance anomalous viscosity and vice versa. Such factors will be those, which, if active to an increased degree, give rise to precipitation or gelling of the solution, like addition of a non solvent or of a salt

---

[1] Departure from NEWTONian flow is often termed "structural viscosity". According to J. DUCLAUX (*Viscosité*, Paris 1934, p. 40), this terminology is physically incorrect and should not be maintained.

[2] Cf. E. HAKSCHEK, *The Viscosity of Liquids*, London 1928. For a short survey of the subject cf. J. DUCLAUX, *Viscosité*, Paris 1934.

[3] Cf. S. A. GLÜCKMANN, *Acta physicochim.* URSS, 13 (1940) 379; P. H. HERMANS, J. J. HERMANS, and D. VERMAAS, *Kolloid-Z.*, 105 (1943) 199. The considerations given here refer to "static" deformation. For *very rapid* changes of the shearing tension, falling within the order of magnitude of the relaxation time of the temporary junction points, solutions of high molecular substances may actually behave as a solid gel showing a modulus of rigidity. We may refer here to the interesting studies of J. D. FERRY, *J. Am. Chem. Soc.*, 64 (1940), 1323 1330; *Ann. New York Acad. Sci.*, 44 (1943) 313, who subjected such solutions to rapid transverse vibrations and found that they then actually behave as a solid with a relatively high modulus of rigidity. A detailed analysis of the mechanism of these elastic deformations was offered by this author.

(which, in greater quantities, cause demixture), a temperature change towards the gelling temperature of the solution etc. Such factors will increase the number of junction points and lengthen their average life-time. The junction point spectrum is broadened and shifted to higher cohesive energies.

As soon as a sufficient part of the junction points have reached a life-time which is practically infinite, a permanent network structure is formed and the system will acquire a property characteristic of the solid state: it shows a yield value. It can resist flow up to a given shearing tension and behaves as an elastic solid below that tension. The magnitude of the yield value will depend on the number and the energy of the permanent junction points.

If the yield value is high enough to be easily detected without very sensitive instruments, we say that the system is a gel. It will, however, be clear that the transitions between the liquid and the gel state are gradual ones, since the yield value will continuously rise from zero to greater values. This is a characteristic feature of any sol-gel transformation.

Actual phenomena are, moreover, complicated by time factors which may be embodied in the term hysteresis. If the conditions in a macromolecular solution are suddenly changed (e.g., by addition of a non solvent or by a change of temperature) a certain time elapses before the new equilibrium condition is reached. This is also illustrated by recent experiments on the influence of ultrasonics on the viscosity of solutions of gelatin in water, rubber in toluene etc. If such solutions are subjected to the action of an ultrasonic field of moderate intensity, their viscosity measured at once after the treatment is markedly less than before, but, after a certain lapse of time, reassumes its initial value (or a value very near to it)[1]. The junction points are temporarily loosened due to "microscopical agitation" arising from the action of ultrasonic waves, but restore themselves to equilibrium condition at low rates of flow after a certain time.

Fig. 9.   Apparatus of MICHAUD for examining the yield value of solutions.

The shearing tension to which a system will yield depends not only upon the character of the junction points but also on their number per unit of volume. Even junction points of a relatively high strength might be disrupted by a relatively small flow gradient imposed on the liquid, if their number per unit volume is small. If the solution of a substance which can set to a gel is more and more diluted, the yield point will become smaller and smaller and beyond a certain degree of dilution the system may assume a condition intrinsically corresponding to the gel state, but with the outward appearance of a liquid.

    [1] Cf., e.g., H. FREUNDLICH, and D. W. GILLINGS, *J. phys. Chem.*, **41** (1937) 1151; *Trans. Faraday Soc.*, **34** (1938) 649; **35** (1939) B 19; S. ONO, *Rev. Phys. Chem. Japan*, **14** (1940) 101; K. SOLLNER, *J. phys. Chem.*, **42** (1938) 1071.

In § 3a we have already referred to the case of very diluted gelatin solutions. At temperatures below the gelling temperature of more concentrated solutions, the liquid appearance of the system is maintained. Yet with sensitive means it can be shown that the liquid is elastic, e. i., that it exhibits a yield value[1]. A simple but very sensitive method to examine such low yield values, we owe to MICHAUD[2]. The apparatus used by this investigator consist of two flasks connected by the horizontal capillary (Fig. 9). The level of the liquid in both flasks being equal, there is no flow in the capillary. Now the liquid level in one of the flasks is adjusted slightly over that in the other flask by lowering a glass rod (A). The liquid in the capillary is observed in the microscope (to render its movements visible a small quantity of carbon black is added to the liquid). If the solution behaves like a real liquid, it will flow through the capillary until the levels in both flasks are again equal. If there is a yield point a small difference of both levels will not cause the solution in the capillary to flow; the carbon particles are only seen to move away a little from their original position to assume a new equilibrium position. The solution resists a certain hydrostatic pressure. From the maximum possible displacement of the particles in the axis of the capillary the "rigidity" (yield value) of the solution μ can be calculated in c. g. s. units (g/cm^2) from the equation

$$\mu = pr^2 / 4lx$$

where p is the hydrostatic pressure, r and l the radius and the length of the capillary and x the displacement of the particles.

Pure liquids never show a yield value, only certain solutions do.

The following table shows the yield point of dilute agar-agar solutions as measured by MICHAUD:

TABLE 1

YIELD VALUE OF AGAR-AGAR SOLUTIONS MEASURED BY MICHAUD

concentration	yield value g/cm^2
0.05	0.004
0.06	0.036
0.08	0.32
0.10	1.23
0.20	27
0.40	525

It will be seen that the yield value shows an extremely rapid rise with the concentration and it is clear that at still higher concentration the system has the appearance of a gel in the common sense. At temperatures above the "melting point" of the more concentrated gels, neither dilute nor concentrated solutions show a yieldpoint. Non gelling solutions, as e. g., solutions of cellulose nitrate in acetone, do not show a yield-point even at high concentrations. They flow, though — owing to their high viscosity — very slowly, even at the smallest shear. They do show, however, non-NEWTON-

[1] SCHWEDOFF, *J. de Phys.*, 8 (1889) 341; 9 (1890) 341.
[2] MICHAUD, *Ann. de Phys.*, 19 (1923) 63.

ian flow as a result of the presence of "short living" junction points (association). The absence of a yield point, however, proves that no permanent junction points between the molecular chains occur. Solutions of this kind may be called true solutions, whereas the solutions showing a yield value might be called "gel-solutions". The latter should be regarded as being thixotropic systems. Their gel structure can be destroyed by mechanical action and they can be forced to flow like a liquid. If the mechanical action ceases, their rigidity is restored (cf. also § 5b).

If gel solutions are exposed to ultrasonic waves their viscosity and elasticity, when measured shortly after the treatment, shows marked changes indicating a loosening of a portion of the junction points[1].

Gel solutions show no well defined viscosity coefficient. Their viscosity depends on the previous treatment. If such a solution has been subjected to large shearing forces, for instance by pressing it through a capillary tube at great speed, its fluidity, when measured at once after the experiment, has changed. After some time, however, the solution returns to its former condition[2]. This phenomenon, narrowly related with thixotropy, is doubtless due to a mechanical disturbance of a structure, a portion of the junction points corresponding to the equilibrium condition being disrupted by the shearing forces. Upon rapid cooling or heating, such solutions also show a marked hysteresis[2].

Some very interesting phenomena have been observed by OUWELTJES[3] in the liquefaction of gelatin gels by a rise in temperature. He prepared a series of gels with increasing concentration. Small conical brass weights were placed on the top of the gels, the point pressed into the gel. The test tubes containing the gel were then warmed up in a water bath at a constant rate of 4—5° per min. The gels melt rather abruptly under these conditions and the moment that the top of the conical weight sinks under the gel surface was considered as the melting point. Furthermore the time elapsed until the weight reached the bottom of the test tube was measured. It represents a measure of the viscosity of the gel just above its melting point.

In Table 2 some observations are collected. It is seen that the melting point raises with the concentration.

TABLE 2

MELTING POINT AND VISCOSITY JUST ABOVE THE MELTING POINT OF GELATIN GELS ACCORDING TO OUWELTJES

conc. of the gel	melting point	falling time
1.5	24.3	392
2	26.6	231
3	28.5	128
5	29.9	78
10	31.3	36
15	32.3	44
20	33.2	50
25	34.1	61

The falling time very rapidly decreases at first, then reaches a minimum value to increase again at still higher concentrations. Since the viscosity of the solutions increases with the concentration,

[1] E. HEYMANN, Trans. Faraday Soc., 31 (1935) 846. In this paper other interesting facts on hysteresis in methylcellulose solutions in water are described to which we may refer here.

[2] Cf. e. g., H. GARRETT, Phil. Mag. (6), 6 (1903) 374; E. HATSCHEK, Kolloid-Z., 13 (1913) 88; Proc. Phys. Soc. London, 28 (1916) 274.

[3] J. L. OUWELTJES, Thesis, Amsterdam 1942.

the falling time should be expected to increase continuously. Obviously, something particular is occurring at the lower concentrations, since the lower temperature during the determination cannot account for the tremendously increased falling time. (For a 15% gel whose melting point was lowered to 28° by the addition of potassium rhodanide, a falling time of only 70 sec was observed). OUWELTJES' explanation is as follows. Diluted gels have a "composite" structure. They consist of small gelled regions with a constitution comparable to that of the more concentrated gels. These aggregates are connected to each other at their periphery by junction points between protruding molecular fringes. Whereas the junction points in the primary aggregates are frequently connected by common chains, this less frequently occurs between junction points belonging to different primary aggregates. This idea is diagrammatically illustrated in Fig. 10.

Just above the melting point of the diluted gels, larger aggregates still exist in the solution and then will give rise to a temporarily high viscosity. In other words, the coherence of the diluted gels is already destroyed before the entire spectrum of junction points occurring in the system is loosened, whereas in more concentrated gels this condition is reached not before practically all the junction points are "molten". This is also consistent with the lower melting point of the diluted gels.

Fig. 10. Diagrammatic representation of crystallites.
a) connected by common chains,
b) connected by junction points between chain fringes.

It will be clear that we are dealing here with very complex and by no means fully elucidated phenomena. Nevertheless, the modern ideas on gel structure and gel formation give a plausible qualitative interpretation and lift the veil of mystery which, until recently, was spread over the subject.

g. Electrical conductivity and diffusion

The electrical conductivity of systems containing electrolytes does not show any appreciable change at the sol-gel transformation, despite the tremendous change in viscosity. This proves that the mobility of the ions is not hampered and, hence, that the gel must contain an anastomising system of capillary spaces filled with the unchanged solvent. The "micro-viscosity" of the system has remained unchanged.

The same conclusion is reached from diffusion experiments. The velocity of diffusion of low molecular substances in a gel is of the same order of magnitude as that in the pure solvent (Cf. § 7a). The idea, often suggested in the past, that considerable quantities of solvent would be fixed in the gels (solvation and hydration theories) is erroneous. True, a certain amount of solvent may be bound either intramicellarly in the crystalline junction points or on the molecular fringes of the amorphous portion, but this quantity is very limited and seldomly surpasses layers of one molecule (or of a few molecules at most in exceptional cases)[1].

In particular cases diffusion in gels may give rise to unusually sharp boundary lines visible with the microscope or even to the naked eye. This has been observed in the diffusion of a swelling

[1] Cf. A. DOBRY, J. chim. phys., 35 (1938) 387; 36 (1939) 296.

agent into xerogels (as e.g., water in dry cellulose [1] or organic solvents in cellulose acetate [2].) This will always occur if the velocity of diffusion is a very steep function of the degree of swelling of the gel. Sharp boundary lines are also observed if a low molecular compound or ion diffuses in a swollen gel and the diffusing molecule or ion is capable of entering into a chemical combination with the gel substance by which it is bound and immobilized (e. g., in the diffusion of zinc or copper ions into a cellulose xanthate gel [3]). For the theory of this phenomenon we must refer to the literature cited.

§ 5. MECHANICAL EFFECTS ON THE SOL-GEL TRANSFORMATION

a. Mechanical disturbance of gel formation

In Section 4 f (p. 504) the influence of mechanical agitation on the flow properties of so called "gel solutions" has already been discussed. Solutions of this kind may be regarded as gels with a small number of junction points per unit volume. Upon mechanical agitation of the liquid, the junction points may be disrupted, but restore themselves when the system is kept at rest. In more concentrated solutions, e. g., in solutions of gelatin rapidly cooled below the sol-gel transition temperature, where the junction points are still weak ones, similar phenomena may be observed if the experiment is carried out soon after the cooling. If more time has elapsed, the equilibrium state, the solid gel, is reached. By mechanical means it can then, of course, also be broken and divided into smaller fragments which, however, do not reunite to a coherent gel upon standing.

BUNGENBERG DE JONG and coworkers [4] have described experiments with solutions of agar-agar and gelatin, which were kept in vigorous movement during gelation by placing the warm solutions in bottles on a shaking machine and allowing them to cool. The solutions then remained liquid and obviously represent a suspension of very small gel fragments. Upon standing they showed a slow sedimentation. After a considerable time (∞ 500 h) a sedimentation layer of agar-agar, occupying about 20% of the original volume, was formed. (The original solution contained 0.2% agar-agar and would, when cooled at rest, have yielded a coherent gel).

The question at once arises why the gel fragments obtained in this experiment do not reunite to a coherent gel, when mechanical agitation is over. The explanation is perhaps as follows. The equilibrium condition at the given temperature and concentration of the gel corresponds to a certain spectrum of junction points. Let us say that a definite percentage of the monomeric residues should belong to a junction point and the others be free (i. e., dissolved). Now, whether the gelation takes place at rest or during agitation will make no difference, either in the final spectrum or in the number of junction points, but another kind of structure may be formed in the two cases.

Let us assume that the process begins with the formation of a number of centres of crystallisation (or coagulation). If the liquid is at rest, the centres grow together to form a regular network structure; newly formed weak junction points may continue to grow out to stronger ones. If the liquid is agitated, newly formed junction points which would otherwise unite different centres, may be broken and this will occur

[1] P. H. HERMANS and D. VERMAAS, *J. Polymer Sci.*, 1 (1946), 149.
[2] G. S. HARTLEY, *Trans. Farad. Soc., Gen. Discussion on Swelling and Shrinking*, London 1946; C. ROBINSON, *ibid.*
[3] J. J. HERMANS, *J. Colloid Sci.*, 2 (1947) 387.
[4] H. G. BUNGENBERG DE JONG, and W. A. L. DEKKER, *Biochem. Z.*, 251 (1932) 105.

the sooner the more the latter are remote from each other. The structure is thereby divided in smaller fragments of a more compact structure. Equilibrium condition being attained, the tendency to form new junction points is much lessened and moreover, the chance of molecular chains meeting each other under favourable conditions has now become much smaller. If still some junction points should be formed between the flocculae (which will no doubt be the case) their number per unit of volume will be so small that it is very difficult to detect them experimentally.

Thus it is conceivable that the mechanical agitation helps the normal tendency of the precipitated substance to attain the densest possible state under the given conditions, whereas at rest the frequent interlinking of all the molecular chains yields a more voluminous permanent structure. In view of what has been said about syneresis, we might also say that the agitation brings about a very rapid "syneresis"[1] of the gel lumps.

b. Thixotropy

Some instances of thixotropy have already been met with in the foregoing. Thixotropy is the faculty of a gel to become liquid upon mechanical agitation and to solidify again when agitation is over. It will be clear that conditions favouring thixotropy will be:

1. there shall be a number of junction points per unit volume great enough to allow of the formation of a coherent structure exhibiting a yield value high enough to speak of a solid gel in the ordinary sense.

2. the cohesive energy of the junction points shall be small enough to be easily disrupted by a shearing force applied to the gel.

This implies that a "narrow" junction point spectrum is required. Gels in which the spectrum is broad, (ranging from lower to higher energies of cohesion) will be devoid of thixotropy.

If, in the former case the coherence of the structure is broken by mechanical agitation, the majority of the junction points will give away and may be restored at rest. On the contrary, in the latter case, if the coherence of the gel is broken, only a small proportion of the (weaker) junctions points is broken; the system breaks into larger fragments, which, for reasons outlined in the preceeding section, do not reunite to form a structure with the original yield value.

Since macromolecular systems will generally tend to form broad spectra of junction points, thixotropy will only occur under special conditions giving rise to a narrow spectrum, as e. g., at great dilution or in the initial stages of gel formation, when the first chain to chain contacts are being formed and no "recrystallisation" has, as yet, occurred to an appreciable degree.

As a matter of fact, the most typical examples of thixotropy have been met with and studied in non macromolecular colloid systems. The first example studied concerned thixotropic ferric hydroxide sols[2]. A great deal of work has been devoted

---

[1] Cf. § 6b. 4 p. 573.
[2] COTTON and MOUTON, *Ann. Chim. Phys.* (8), 11 (1907) 186.

to these systems. We shall confine ourselves to the discussion of some general features, since various surveys and a monograph have appeared on the subject elsewhere[1].

Recently, the definition of thixotropy has been made more general, it now being applied to any "isothermal reversible decrease of viscosity with increase of rate of shear". Hence, some phenomena of anomalous viscosity have been included which, as a matter of fact, are related ones, as was especially commented by GOODEVE,[2] from whose paper we shall borrow a part of the following.

First it should be mentioned that thixotropy, like gel formation, is favoured by an anisodiametric shape of the colloidal particles or of the particles of coarser suspensions showing thixotropy. The systems must be capable of forming a continuous cohering structure. Where examples of spherical particles, showing a thixotropic behaviour have been met with, this faculty is always confined to relatively concentrated systems.

Evidently, the particles must be able to form links (junction points) with each other. Many sols become thixotropic under conditions approaching those which lead to permanent coagulation e. g., the addition of carefully dosed quantities of salts to a ferric hydroxide sol. Particles which meet each other should be capable of junction point formation, but the cohesive energy of the contacts shall not exceed a certain limit. Rodlet- and platelet-shaped particles usually show the greatest forces of attraction at their ends or edges[3]. Moreover, MÜLLER[4] has shown that there is an additional statistical factor favouring collisions on the ends of rod-like particles. A partial coagulation of such particles would first take place at the ends. The junction point spectrum of a structure formed by particles of this kind will evidently be a narrow one, since there are no great variations in the way how they meet and mutually cohere. If such a structure is broken by mechanical agitation, it will tend to restore itself to the original condition.

GOODEVE has developed a quantitative theory on the process of continuously breaking and restoring the junction points, which sets in when systems of this kind are subjected to a flow gradient and he came to results consistent with the experimental data on the variation of the viscosity coefficient with the rate of shear (see also the chapter on coiled molecules, p. 106). Thixotropic behaviour (in the general sense as defined above) was shown to depend upon the average time necessary to break the links between the particles under the agency of shear, which in its turn will depend upon the extensibility of the particles. In extensible particles the time required to reach the critical extension at which the link between the particles or the particles themselves will break, is greater than in non extensible ones.

If the average life time of the junction points is small, this quantity will play a part in the flow properties of the system and enters into the mathematical expressions. We have already earlier discussed that the shearing of a continuous structure with junction points having a limited life time, is the base of the usual form of anomalous viscosity in concentrated macromolecular solutions (decrease of viscosity coefficient with increasing shear).

For further details we must refer to the original literature cited.

---

[1] H. FREUNDLICH, *Kapillarchemie*, Leipzig 1932, Vol. II, p. 615; H. FREUNDLICH, *Thixotropy*, Paris 1935; J. DUCLAUX, *Rigidité, Thioxotropie, Coacervation*, Paris 1934, p. 20; G. W. SCOTT BLAIR, *An introduction to Industrial Rheology*, London 1938; cf. *First and Second Report on Viscosity and Plasticity*, Acad. Sci. Amsterdam 1935.

[2] C. F. GOODEVE, *A General Theory of Thixotropy and Viscosity*, Trans. Faraday Soc., 35 (1939) 342.

[3] This was e. g., shown by J. J. KETELAAR for vanadiumpentoxide, *Nature*, 137 (1936) 316.

[4] H. MÜLLER, *Kolloidchem. Beih.*, 27 (1928) 223.

§ 6. SORPTION AND SWELLING

In this section we shall deal with the phenomena connected with the interaction of gels with gases, vapours and liquids. Xerogels, as a rule, have a marked capacity to absorb other substances. If the quantity absorbed is limited and relatively small, we speak of *sorption*. Sometimes the volume of the gel remains practically constant upon sorption (silicagel and benzene), but sometimes its volume increases and sorption is combined with *swelling* (cellulose and water). If the substance absorbed is, in the liquid state, a solvent towards the gel-substance, sorption is nothing but the initial phase of swelling (nitrocellulose and acetone).

For this and other reasons, concerning the mechanism of the processes considered, it is not practical to separate sorption and swelling, since a definite limit can not be traced between the two. Neither can we separate sorption of condensable vapours and swelling in the corresponding liquid substance, since the latter is, for obvious reasons, identical with sorption at the saturation pressure of the vapour.

Sorption of gases and vapours has been considered as being intimately related with the phenomena of adsorption of gases and vapours on solid surfaces. In the sense of classical colloid science gels were depicted as representing disperse systems with a largely developed "inner surface" and one has endeavoured to interpret sorption phenomena as adsorption on this inner surface. Since the sorbent then penetrates throughout the gel the term *ab*sorption is frequently used instead of *ad*sorption. It is however practical to follow Mc Bain's suggestion denoting both processes simply by sorption [1] because it is usually impossible to discriminate between the two, neither from an experimental nor from a theoretical point of view.

Though this classical interpretation may still be preferable in certain cases, it has recently been recognized that in the typical macromolecular gels sorption should be regarded as a homogeneous dissolution of the sorbate in the sorbent quite comparable to the absorption of e. g., carbon bisulphide vapour by paraffin oil or the absorption of water in sulphuric acid.

For systematic reasons we shall first treat the sorption of gases and vapours, thereby including substances like amorphous carbon, though it is doubtful whether the latter belong to the class of gels. After having discussed some theoretical points of view, we shall select some typical systems and shall — in the course of our treatise — gradually find ourselves mixed up with the problems of swelling and dissolution.

a. Sorption

a 1. *Some general remarks on the nature of sorption*

It has already been stated that it is often difficult to tell the phenomena of *ad*sorption from those of *ab*sorption and that *ab*sorption is often nothing but dissolution of the sorbate in the sorbent. This is not surprising when taking into consideration that all these phenomena depend upon the very same kind of molecular forces and may be formally as well as theoretically treated in much the same way.

The phenomena of adsorption depend upon the so-called residual valencies, viz., VAN DER WAALS forces, dispersion forces, dipole forces, which may be summarized as intermolecular cohesion. Just the same forces play a rôle in the phenomena of

---

[1] J. W. Mc Bain, *Philos. Mag.*, 18 (1909) 916; *Z. physik. Chem.*, 68 (1909) 471.

mixing and dissolution. Besides that, the kinetic energy (the diffusion tendency) of the molecules has a bearing on the process of dissolution. The intermolecular forces account for the energy factor and molecular motion for the entropy factor of the process. The "affinity" or free energy change of the process is determined by the sum of the two.

In a sense there is no objection to saying that the absorption of water by concentrated sulphuric acid is due to a surface action of the molecules of the sulphuric acid and hence to the "internal surface" of the latter. The same applies to the absorption of water by solid calcium chloride, whereby a crystalline hydrate is formed. There is only no practical gain in using this terminology.

In the phenomena interesting us here, the only intrinsic difference from the systems mentioned before is the arrangement and the mobility of the molecules of the sorbent and perhaps also the percentage of the molecules of the sorbent, which are in a position to interact with the molecules of the sorbate. In the former instances all molecules of the sorbate are equally apt to interact. In a gel and perhaps to a still greater extent, in systems like amorphous carbon, a certain part of the molecular surfaces may, however, be more or less, or even completely, inhibited from taking part in the process of interaction as a result of their particular structure or of the compactness of the solid sorbent. The laws governing the energetic interaction of the other molecules will, however, be essentially analogous. Considering the equilibrium conditions, the energetic factors will be of exactly similar nature. The entropy factor may, however, require a different treatment taking into account the different arrangement and mobility of the particles of the solid colloid component.

It need not be surprising that e. g., the physical behaviour of the system sulphuric acid-water and the system cellulose-water is strikingly similar as regards vapour pressure equilibrium and thermodynamic character, volume relations and optical behaviour.

In the case of a gel like cellulose swelling in water, or rubber swelling in a hydrocarbon, there is no longer doubt that we are dealing with a real process of dissolution resulting in a homogeneous mixture of both components i. e., with a process of molecular dispersion (Cf. Chapter III). Considering the sorption of a gas like carbon bisulphide vapour by amorphous carbon, a homogenous mixture of the components evidently does not occur. The classical concept of surface adsorption on the extensively developed surface of the solid will then, most probable, represent a more adequate picture of the process.

The physico-chemical aspects of the latter would, nevertheless, from an energetic point of view, have very much in common with those of a homogeneous solution of carbon bisulphide in carbon, if this would exist, since in both cases the energetic interaction would arise from exactly similar intermolecular forces.

Venturing to deviate from the tradition established by the majority of earlier textbooks, we shall endeavour to treat the phenomena of sorption from this point of view making allowance, however, for the process of "capillary condensation" which may interfere as a special additional feature of the sorption by solids not occurring in the general theory of dissolution, though not always, nor necessarily, giving rise to a particular behaviour.

a. 2. Some basic theoretical principles

There is an extensive literature on the theory of sorption and the subject involves a great many aspects which can not be completely dealt with here, since this would

lead us beyond the scope of this volume[1]. We shall endeavour to select and briefly outline some basic theoretical viewpoints which may serve to a general understanding and interpretation of the experimental facts.

a. Molecular interaction and deviation from RAOULT's *equation*[2]. It is a well known fact that the vapour pressure of mixtures of two liquid components 1 and 2 follow RAOULT's equation if there is no energetic interaction between the components (athermal mixture) and mixing is entirely due to the diffusion tendency (provided the entropy of mixing is "ideal"). The partial vapour pressure of the components are then given by

$$p_1/P_1 = x_1 \text{ and } p_2/P_2 = x_2$$

where P_1 and P_2 are the vapour pressures of the pure components, and x_1 and x_2 the respective molar fractions. The vapour pressure of the mixture is additively composed of the partial pressures of the components (Fig. 11A).

Deviations from this behaviour occur if heat is evolved or absorbed on mixing,
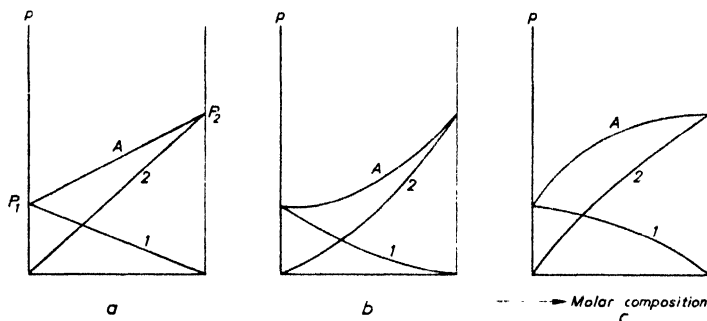
Fig. 11. Vapour pressure curves of binary mixtures.
a. ideal behaviour, b and c positive and negative deviations from RAOULTS' law.
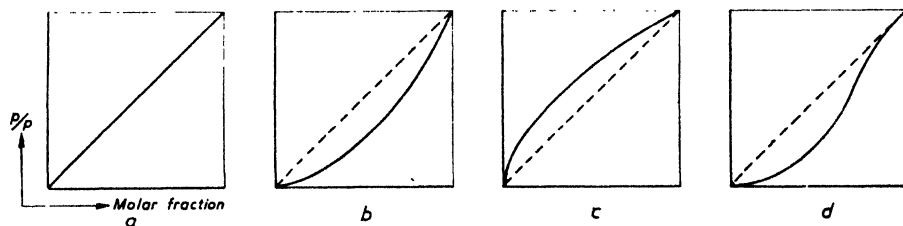
Fig. 12. Vapour pressure curves according to Fig. 11, if one component is non-volatile.
Fig. 12d. Scheme of vapour pressure curve as, e. g., occurring in the system sulphuric acid-water.

---

[1] For a comprehensive treatment of the theory of sorption cf. E. HÜCKEL, *Adsorption und Kapillarkondensation*, Leipzig 1928; O. GLÜH and N. STORK, *Die Adsorption*, Braunschweich 1928; A. EUCKEN and E. JAQUET, *Theorie der Adsorption von Gasen und Dämpfen*, Berlin 1925; J. W. MC BAIN, *The sorption of gases by solids*, London 1932; H. FREUNDLICH, *Kapillarchemie*, Vol. I, Leipzig 1930; and particularly: S. BRUNAUER, *The Adsorption of Gases and Vapours*, London 1943.
[2] Cf. HILDEBRAND and coworkers, numerous papers in the *J. Am. Chem. Soc.*, 1916—1923; concise survey: S. GLADSTONE, *Recent advances on physical chemistry*, London 1931, p. 349—383.

or if the entropy of mixing is not ideal. In low molecular substances the former factor is usually the principal one Generally we may say that deviations from RAOULT's law will be found if the "affinity" of the process is greater or smaller than that of an ideal athermal mixture. Fig. 11b and 11c give the vapour pressure for the cases that the affinity is greater or smaller respectively. In the former case the mixture will usually be exothermic and in the latter case endothermic, dissolution being entirely due to the entropy of mixing which exceeds the absorption of heat.

If now one of the components, e. g., no. 1, is non-volatile, the vapour pressure curves of the mixtures in the three cases considered will be represented by the schemes shown in Fig. 12a, b and c, where now the relative vapour pressure p/P of the volatile component is plotted against the molar fraction of the latter.

It is seen that in the case of large affinity (positive heat effect, preferred mutual attraction between the component molecules) the vapour pressure curve is concave to the pressure axis, whereas in the case of sub-ideal affinity (negative heat effect, cohesive energy of the same species larger than that between alien ones) the curve is convex to the pressure axis.

Fig. 12d shows a particular case of the type 2b. Here the mixture shows ideal behaviour at great dilution of the non-volatile component, whereas at low concentrations of the volatile component the behaviour of an exothermic mixture is clearly demonstrated. Such an S-shaped curve is e. g., observed in the system sulphuric acid-water, where the formation of strongly exothermic compounds (hydrates) of the acid is known to exist. The occurrence of such specific chemical binding forces will always tend to result in a vapour tension curve which is concave to the pressure axis in the region where compound formation dominates the process of mixing. It would seem that in this case the mixing of the hydrates with further amounts of water is a less thermally positive or even a thermally neutral process.

In Fig. 12 the composition is expressed in molar fractions:

$$x = \frac{n}{n_o + n}$$

(n_o moles of the non volatile component mixed with n moles of the volatile one). In absorption experiments the composition is usually expressed as the weight of the volatile component absorbed per unit weight of the sorbent n/n_o. In the ideal case (Fig. 12a) we have:

$$\frac{n}{n_o + n} = \frac{p}{P} \quad \text{and hence} \quad \frac{n}{n_o} = \frac{p/P}{1 - p/P}$$

The second equation represents the "absorption isotherm" of the system. It is a hyperbola of the shape shown in Fig. 13 which is convex to the pressure axis. The vapour pressure P of the pure volatile component is approached at $n/n_o = \infty$.

It will be clear that all systems having a negative heat of absorption (generally those having a smaller affinity than the ideal system), will also yield sorption isotherms convex to the pressure axis. Curves concave to the pressure axis hence indicate an affinity sensibly larger than that of the ideal system.

β. *The Langmuir isotherm.* LANGMUIR has considered the adsorption of molecules

from a gaseous phase on a flat solid surface, thereby considering the case that the
binding of the former occurs at discrete localized centres of attraction (active spots)
on the surface, such as e. g., might be expected when considering adsorption on the
surface of a crystal, where the surface will be covered by a regular pattern of residual
affinities (cf. schema in Fig. 14).
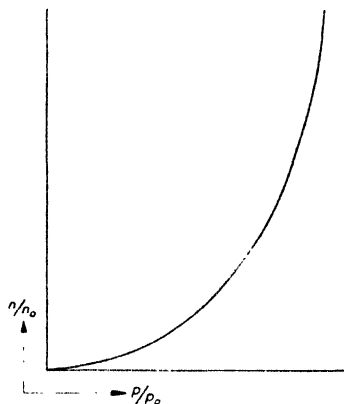
Fig. 13. Absorption isotherm
of the ideal system in Fig. 12.
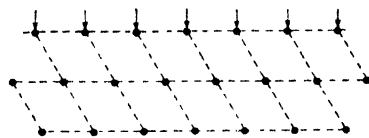
Fig. 14. Crystalline adsorbing surface.

Considering an equilibrium state and denoting the total number of centres
of attraction per unit area by N, and those occupied by a gas molecule by n, the
number of gas molecules bound on the remaining $N - n$ free spots per unit of time
will be proportional to $N - n$ and to the pressure of the gas p. The number of molecules
leaving the surface will be simply proportional to n and hence the condition of equili-
brium will be

$$\beta(N - n)p = n\alpha$$

where α and β are constants. Hence.

$$\frac{n}{N} = \frac{\beta p}{\alpha + \beta p} \tag{1}$$

Setting $\beta/\alpha = b$ and substituting n/N by a/a_s where a and a_s represent the
amount of adsorbed gas per g of the sorbent at the pressure p and at saturation
$(n/N = 1)$ respectively[1], we obtain

$$a = \frac{a_s\, bp}{1 + bp} \tag{2}$$

representing the LANGMUIR equation. Similar equations have been given by J. PERRIN
and J. FRENKEL[1].

The theory embodied in equation (2) implies two conditions:
1. Localized binding spots of equal energy on the surface of the solid acting inde-
   pendently of each other, i. e., the individual binding energy is independent of
   whether neighbouring spots are occupied or free.

---

[1] In this formulation occupation of all places N is reached at the limit of infinite pressure only.

2. Adsorption of at most a monolayer of gas molecules on the surface of the solid, i. e., if a spot is occupied it exerts no further attraction on other molecules.

Equation (1) may be derived in a more explicit and exact manner introducing the individual binding energy ε and taking into consideration the entropy of the system[2]. It then takes the form

$$\frac{n}{N} = \frac{cpe^{\varepsilon/RT}}{1 + cpe^{\varepsilon/RT}} \tag{3}$$

where c is a constant which is a function of T only, i. e., a constant at a given temperature, also appearing in the equation expressing the thermodynamical potential Z of the gas phase:

$$Z = RT \ln c + RT \ln p$$

---

[1] Cf. D. LANGMUIR, *J. Amer. Chem. Soc.*, 38 (1916) 2221; 39 (1917) 1848; 40 (1918) 1361; J. PERRIN, *J. chim. phys.*, 28 (1923) 508; J. FRENKEL, *Z. Physik*, 26 (1926) 117.

[2] Cf. P. H. HERMANS, *Contribution to the physics of cellulose fibres*, Amsterdam 1946, Appendix I by J. J. HERMANS. Compare also E. HÜCKEL, *Adsorption and Kapillarkondensation*, Leipzig 1928, p. 159. We shall briefly indicate here how equation (3) may be derived.

The general condition of equilibrium between the solid surface and the gas phase is

$$\frac{\partial Z_s}{\partial n} = Z_g = Z_0 + RT \ln p \tag{4}$$

where Z_s and Z_g represent the GIBBS potentials for the two phases and Z_0 is a function of T only. Since any mutual interaction of the molecules absorbed is presumed to be absent, the energy E of the surface will simply be a linear function of the number of absorbed molecules n:

$$E = E_0 - \varepsilon n \tag{5}$$

The total number of available and equivalent places being N, the entropy of the system can be calculated as follows: The number of ways in which the n molecules may be distributed on the surface is:

$$W = \frac{N!}{n! \, (N - n)!}$$

and the entropy S will be

$$S = R \ln W = -R \, [n \ln n + (N - n) \ln (N - n)] \tag{6}$$

Neglecting the term pV in the solid we obtain

$$Z_s = E - TS = E_0 - \varepsilon n + TS, \text{ and hence}$$

$$\frac{dZ_s}{dn} = -\varepsilon + RT \ln \frac{n}{N - n} \tag{7}$$

From (4) and (7), setting exp. $Z_0/RT = c$, equation (3) is easily derived.

The LANGMUIR isotherm is represented by a hyperbola of the type schematically shown in fig. 15. (Quantity of sorbate per g (or per unit surface) of the sorbent plotted against the pressure). The curvature is concave to the pressure axis in contrast to Fig. 13. This is in conformity with the positive heat effect assumed.
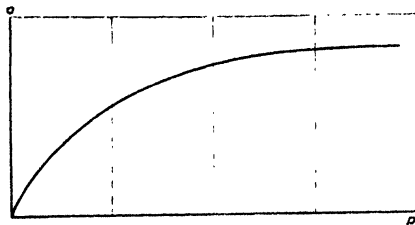
Fig. 15. LANGMUIR-type isotherm.

If we consider an amorphous solid having centres of attraction irregularly distributed on its surface, the same reasoning will invariably hold, provided the two conditions mentioned above be still fulfilled[1].

The mathematical treatment set forth here is not confined to the case of surface adsorption. J. J. HERMANS[2] has stressed that exactly the same reasoning may be applied to the equilibrium of the gas phase with a solid structure containing throughout its body a number of N mutually independent active spots per unit volume, in which the molecules of the gas phase may penetrate by diffusion and where they can be bound according to the conditions 1 and 2 (p. 516). Such conditions may very well apply to certain instances of adsorption of gases or vapours in gel-like systems (cf. Section 6a 3 γ.) giving rise to "solid solutions" which should theoretically be considered as homogeneous binary systems[3].

The question now arises as to how the curves represented by equation (3) change if various values of the characteristic constants are substituted. It can be easily shown that we may write

$$ce^{-f/RT} = ce^{(f_g - f)/RT}$$

where $f_g - f$ represents the free energy difference of a gas molecule in the gaseous and in the adsorbed state respectively for a given standard condition ($p = 1$ and $n/N = {}^1/_2$).

The constant b in the LANGMUIR equation (2) hence represents a quantity which we may designate as the "affinity" between sorbent and sorbate.

In Fig. 16 it is shown how the shape of the sorption isotherms corresponding to the LANGMUIR equation changes according as this affinity becomes smaller.

As a rule the curves will vary in the same sense if the binding energy becomes smaller, however, with regard to the second law of thermodynamics it is more correct to consider the affinity as a parameter.

Fig. 16. Curves representing the LANGMUIR equation (3). The parameter $ce^{f/RT}$ decreasing from no. 1 to no. 5.

According as the affinity becomes smaller, the sorption isotherm will more and more approach to a straight line of low inclination.

There is a formal analogy between this result and the vapour pressure isotherms of solutions (section *a*). This becomes clear if we take into consideration that in the relevant systems the molecular weight of the volatile component will always be very small as compared to that of the sorbent. The molecular fraction $n/(n_o + n)$ will then remain small and nearly proportional to n/n_o. For small values of n/n_o the hyperbola of the athermal ideal mixture will also appear as a straight line. The difference which remains between the LANGMUIR isotherm for heats of absorption approaching zero on the one hand and the vapour pressure of the ideal mixture on the other hand is not due to a difference in energetic conditions but to the finite number N of available places where the vapour can be bound, in contrast to the unlimited number of molecules which can be stored at mixing.

---

[1] Only a slight change in the calculation of the entropy of the system may have to be introduced.

[2] Loc. cit.

[3] In this connexion also see the recent paper by A. J. HAILWOOD and S. HORROBIN, *Trans. Faraday Soc.*, 47 B (1946) 84, *Gen. Discussion on Swelling and Shrinking*, held at London (1946).

We may conclude that the two formulations give rise to a nearly straight initial part of the sorption isotherm if the affinity is small, and to a curve concave to the pressure axis if there is a sensible positive heat effect, it being understood that the amount of sorbed substance is expressed in moles (or g) of the sorbate per mole (or g) of the sorbent. A negative heat effect of sorption is physically senseless in the LANGMUIR picture.

It is an essential result of the foregoing, that, particularly in cases of the LANGMUIR type of sorption, it will often be impossible tot discriminate between 2- and 3-dimensional sorption, since both are then equivalent from the kinetic as well as from the thermodynamical viewpoint.

Recalling that in the LANGMUIR equation saturation is reached for $n/N = 1$ and only at infinite pressure, it is clear that this formalism can never represent the behaviour of a system comprising condensable vapours below the critical temperature up to their saturation pressure. Its applicability will then be confined to the lower pressure range.

γ. Complicating factors and some further considerations. The condition 1 underlying LANGMUIR's picture (cf. page 516) is comparable with the assumptions introduced in the theory of dilute solutions (exclusion of energetic interaction between the molecules of the solute). As soon as this condition is no longer fulfilled, similar complications will arise as in the treatment of concentrated solutions, and quantitative theoretical considerations become very difficult.

The LANGMUIR equation was shown to be in perfect conformity with experiments dealing with the adsorption of permanent gases at very low pressures on the surface of mica lamellae cooled to the temperature of liquid air. The same formulation appears to apply, however, more or less well to a great many of other more common instances of sorption of gases and vapours by an amorphous sorbent at less extremely low pressures and at ordinary temperatures[1].

It has been argued already in the preceding sections that a regular spatial pattern of the binding spots is no essential condition. If several species of binding spots of unequal binding energy *ε* exist, it will be clear that still sorption isotherms of a similar type will be found, which may be regarded as arising from a superposition of the separate isotherms of the individual species. In a micro- or krypto-crystalline substance spots of different binding energy will always occur, since it is a well known fact that the magnitude of the residual affinities at the corners and edges of crystals are different from those on a lattice plane. Lattice distortions and amorphous components of the system will also contribute to the spectrum of different "attractive forces".

Complications of another kind will arise if there is an energetic interaction between neighbouring molecules of the sorbate, or if, in the case of 2-dimensional sorption, more than a monolayer of molecules is adsorbed on the surface. Several theories showing features similar to those introduced in the theories of concentrated solutions have been tried. However, no more than in the case of solutions, is it necessary to

---

[1] A. MATTHES, *Kolloid-Z.*, 108 (1944) 79 has shown that the sorption of SO_2 on charcoal is very well represented by the LANGMUIR equation up to a rel. vapour pressure of 0.5 at — 10°. According to papers of L. M. PIDGEON, *Canad. J. Research*, 12 (1935) 41; L. H. REYERSON and A. E. CAMERON, *J. Phys. Chem.*, 39 (1935) 181 the same applies to the sorption isotherms of benzene and alcohol on silica gel and of bromine and iodine on amorphous carbon.

go into the details of this extremely difficult subject in order to reach a general understanding of the phenomena involved.

A special case of interaction between sorbed molecules may, however, be briefly referred to here. In Fig. 17 two cases of 3 dimensional sorption in a solid-gas equilibrium and the corresponding isotherms are schematically represented. Fig 17a corresponds to that of the LANGMUIR equation.

The occupied specific "elementary spaces" (marked by circles) are distributed at random according to the laws of probability and the isotherm is a hyperbola, represented by equation (3).

Fig. 17b corresponds to a case that binding at places lying next to places already occupied are energetically preferred to a very marked degree. Then coherent blocks of occupied places will be formed. The active spots lie as indicated by the arrow in Fig. 17b.[1] In the equilibrium state, the number of molecules bound on the active spots in unit time will now be proportional to the number ν of the latter and to the pressure p and hence, be $\alpha\nu p$ (where α is a constant). The number of molecules escaping from the active spots is $\beta\nu$ (β = a constant). Equilibrium condition is hence $\alpha\nu p = \beta\nu$ or $p = \beta/\alpha = $ constant.

The sorption isotherm is a vertical straight line. As soon as a pressure $= \beta/\alpha$ is reached, it remains constant during the entire process of sorption and cannot be further increased until all available places are occupied. This case is obviously realized in the formation of crystalline compounds of the solid with the vapour phase such as e. g., the hydrate $CaCl_2 . H_2O$ from anhydrous calcium chloride and water-vapour. Crystals of the hydrate grow at the cost of the anhydrous substance. It is well known that step-ladder isotherms are then found.

Whereas in Fig. 17a the binary system has two phases, three phases are present in Fig. 17b and in accordance with the phase rule, the system will be monovariant.

Fig. 17. Two extreme cases of 3-dimensional sorption with the corresponding isotherms. a. LANGMUIR equation b. crystalline compound formation.

Of particular interest in connexion with our subject is the case of *compound formation by a macromolecular solid* e. g., the hydrate formation by cellulose and gelatin or the formation of an addition compound between nitrocellulose and acetone. Whereas in ordinary low molecular hydrates the composition of the successive compounds X, X . H_2O, X . $2H_2O$ etc. differs considerably as regards the percentage of water and, hence, the GIBBS potentials of these compounds differ by considerable jumps, the situation is different in macromolecular substances. If e.g., each monomeric residue R of a molecule consisting of a chain of n residues can bind one water molecule, the following hydrates are possible:

$$nR . H_2O - nR . 2 H_2O - nR . 3 H_2O \ldots\ldots\ldots\ldots nR . n H_2O, \text{ where } n \text{ is a}$$

large figure.

The composition hence, changes quite gradually in small steps according as the water content increases and the GIBBS potential per mole also changes in small

---

[1] The bound molecules already surrounded by other bound molecules are considered as permanently bound with a still higher energy.

steps. It can then easily be shown[1] that the sorption isotherm of the system will be a LANGMUIR isotherm if n is a sufficiently large figure. The latter then may be interpreted as representing the usual step-ladder curve of successive hydrates, as known from low molecular compounds, in the limit of an infinite number of hydrates, such as they will occur in macromolecular hydrates.

Three dimensional sorption in a gel, whereby compound formation plays a rôle, will, consequently, also give rise to isotherms of the LANGMUIR type. The argument sometimes found in earlier literature that the nonexistence of a step-ladder isotherm disproves the formation of true stoichiometric compounds is, hence, by no means a valid one.

Furthermore, attention should be directed to the different course which sorption phenomena will take in higher pressure ranges according as the temperature of observation is below or beyond the critical temperature of the gaseous component. If 2-dimensional sorption is concerned and the temperature is below the critical temperature of the vapour, condensation will occur when the saturation pressure is approached. A liquid layer will be deposited on the surface and the isotherm will suddenly rise, the quantity of liquid on the surface increasing while the pressure remains constant. (See Fig. 18).
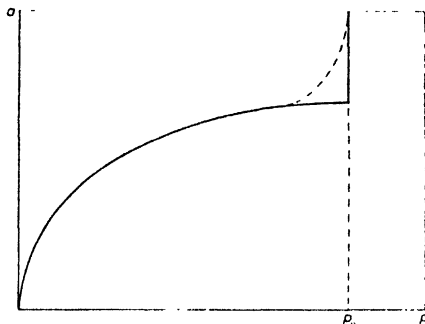
Fig. 18. Scheme of sorption isotherm continued beyond the saturation pressure p_o of the vapour.

Instead of a sudden change in the course of the isotherm (full drawn curve) the transition to condensation may be a more or less gradual one (dotted line) since very thin liquid layers of a few 2 or 3) molecules thick have a lower vapour pressure than those of several molecules thick. Theoretical considerations[2] show that, ordinarily, such anomalous transition layers will be hardly thicker than 2 molecules. If however polar molecules are concerned considerably thicker layers may occur, a result to which we will refer below. In this instance considerable quantities of liquid may condense before the saturation pressure is reached.

The surface tension at the solid-liquid interface may also have a bearing on a less sudden change of direction of the isotherms near the saturation pressure. The solid surface may be more or less readily wetted by the liquid.

A rapid rise of a sorption isotherm with a curvature convex to the pressure axis will often, though not necessarily, indicate that some sort of condensation process is taking place. In as far threedimensional sorption in a gel is concerned, we shall later on meet cases where another explanation is to be preferred. The vapour pressure isotherm of the system sulphuric acid-water, dealt with in Fig. 12d, where a similar change of curvature occurs, obviously shows that also quite other factors than condensation may give rise to the same behaviour.

Returning to surface condensation, it is to be recalled that the curvature of the surface will have a bearing on the pressure at which condensation occurs. Hence a perfectly flat surface and a rough one will behave differently. If we are dealing with a "porous" sorbent containing minute capillaries, the inner surface of the latter represents a surface with a curvature. It is a well known fact that the vapour pressure of a liquid in a capillary, wetting the walls of the latter, shows a vapour pressure depression depending on the diameter of the capillary. Studying 3-dimensional sorption in porous bodies (and gels may to a certain extent be also regarded as such) we may, therefore, towards

---

[1] Cf. J. J. HERMANS, footnote 2, p. 517
[2] E. HÜCKEL, Adsorption und Kapillar Kondensation, Leipzig 1928, p. 215.

higher pressures expect a transition of sorption phenomena of the kind dealt with in the precedent sections to those due to "capillary condensation".

Finally a mechanism should be mentioned, which is probably an essential one in many instances of sorption. BRUNAUER, EMMETT, and TELLER[1] considered the case of a surface on which there are a number of sites with a specific affinity for the sorbent. They considered the case that the formation of a layer of loosely bound molecules of the sorbate starts to form before all the specific sites on the surface are occupied. It is then assumed that the binding energy of these loosely bound sorbate molecules is not different from that in the liquid sorbate itself. Applying a similar kinetic reasoning as LANGMUIR used in deriving his sorption equation, the authors arrived at a formula which was consistent with experiments in a number of cases. Later CASSIE[2] gave a more exact thermodynamical treatment of the case leading to the same formula:

$$\frac{p}{a(p_o - p)} = \frac{\beta}{B} + \frac{1}{B}(1 - \beta)\frac{p_o}{p}$$

where a is the total amount of sorbate, B the number of active sites and a a constant characteristic of the sorbate-sorbent system. Plotting the left hand part of the equation (which contains experimental data only) against p/p_o a straight line is obtained.

As CASSIE showed, the essential point is, that there is an entropy of mixing between the sorbate-molecules bound at the specific sites and the other loosely bound ones. This gives rise to an "affinity" for the latter (cf. page 515). It will be evident that this case is not confined to 2-dimensional sorption but equally applies to 3-dimensional sorption, and solves the apparant contradiction that sorbate molecules bound by ordinary liquid-liquid forces are in equilibrium with a vapour of vapour pressure lower than p_o.

Capillary condensation[3]. In a porous body containing capillary spaces communicating with the outer world, adsorption on the surfaces of the walls of the capillaries may, at a sufficiently high pressure, be followed by a condensation of liquid in the capillaries.

The well known formula giving the vapour tension p of a liquid in a cylindrical capillary with a diameter r, the surface tension of the liquid being σ, is

$$RT \ln \frac{p}{p_o} = -v \frac{2\sigma}{r} \tag{8}$$

where p_o is the tension of a flat liquid surface and v the molar volume of the liquid.[4]

If the diameter of the capillaries in the sorbent is non uniform and varies continuously, first the finest capillaries will be filled and then successively the coarser ones according as the vapour pressure is increased.

It cannot be exactly said down to which capillary diameter equation (8) holds, but, in view of molecular dimensions, a radius of 20 Å seems to be a lower limit, below which the concept of capillary condensation will loose its physical sense[5]. The value of p/p_o arising from equation (8) is then 0.35 for water at room temperature.

[1] S. BRUNAUER, P. H. EMMETT, and E. TELLER, *J. Am. Chem. Soc.*, 60 (1938), 309.

[2] A. B. D. CASSIE, *Trans. Farad. Soc.*, 41 (1945) 50.

[3] From a historical point of view the researches of J. M. VAN BEMMELEN, *Z. anorg. Chem.*, 13 (1897) 233; 18 (1898) 98; 59 (1908) 255; 62 (1909) 1. *Die Asorption*, Dresden 1910 ana those of R. ZSIGMONDY, *Z. anorg. Chem.*, 71 (1911) 356 should be referred to. For the theory cf. the book of E. HÜCKEL (*loc. cit.* p. 521).

[4] r is given in dynes/cm; R (molar gas constant) $= 8.10^7$ ergs per degree. (Complete wetting is assumed, otherwise the value of r must be corrected, multiplying it with the cosine of the angle of contact with the wall).

[5] This lower limit in the order of $r = 10$ Å is in conformity with the work of M. B. COELINGH, which will be mentioned later (p. 531).

Capillary condensation hence, may be reasonably discussed only at relative pressures higher than that limit.

To values of r, derived by applying equation (8), the thickness of the transition layer (cf. Section ;·) should be added. Assuming but a monomolecular layer, the order of magnitude of this correction is about 2 to 3 Å, hence say about 5 Å for the capillary diameter.

It will be clear that no distinct limit can be traced between the region of adsorption and that of capillary condensation, since there will usually be a gradual transition between the two. Further, if the pores in a system become so minute that their dimensions approach molecular sizes, the degree of dispersion of the sorbate and that of the sorbent will both approach molecular dispersion. Then gradual transitions between dissolution and capillary condensation become apparent.

Attention should be directed to a notable phenomenon which may occur if capillary condensation is involved in sorption. If a capillary is exposed to vapour of increasing tension, the latter will be first absorbed on its walls. When the pressure has reached a value where condensation can set in, the capillary wall will be covered with a liquid layer, (Fig. 19a) whose radius of curvature is only half the radius of
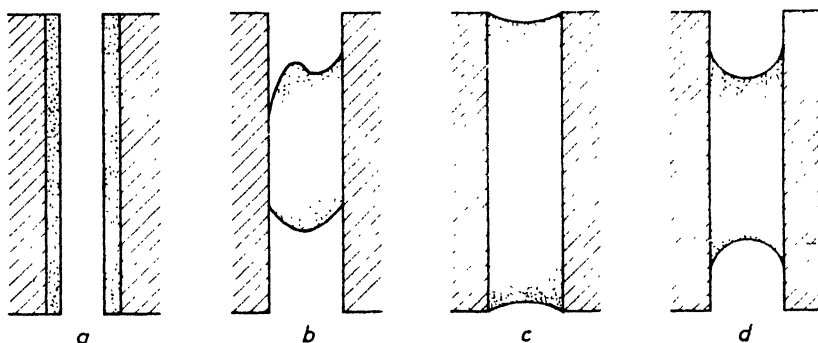
Fig. 19. Diagrams of capillary condensation

the capillary wall[1]. A lower average radius of curvature of the liquid surface may, alternatively, also ensue from incomplete wetting of the capillary walls (Fig. 19b). If at a certain pressure the capillary is completely filled (Fig. 19c), and if the vapour tension is now gradually decreased, the situation becomes as shown in Fig. 19d. The curvature of the liquid surface is then more pronounced than in Fig. 19 a or b. The occurrence of a given quantity of liquid in the capillary will, hence, be observed at a lower pressure. In other words, the vapour pressure isotherm of desorption will be shifted to the left as compared to that recorded upon absorption.

As to the heat effect of capillary condensation, theoretical deductions lead one to expect that it must be positive, though generally very small. This means that at condensation of a liquid in a capillary, the evolution of heat is slightly larger than corresponds to the latent heat of evaporation of the liquid from a flat surface under the conditions of the experiment.

---

[1] M. B. COELINGH, *Thesis*, Utrecht 1938.

As was shown by E. Hückel [1] the intermolecular forces between sorbent and sorbate have a negligible influence on the phenomena of capillary condensation, and their temperature dependence. Effects of the specific characteristics of the materials used are, therefore, not to be expected in this instance. In this respect analogy with the behaviour of solutions ceases in this region from a theoretical point of view. Nevertheless, as we shall see later, instances are met with, where it is difficult to discriminate as to whether effects of capillary condensation or solution are responsible for certain phenomena.

f The FREUNDLICH *isotherm*

For historical and practical reasons mention must be made of another well known formulation often used to cast the outcome of sorption experiments into a mathematical form. This formula, introduced by H. FREUNDLICH, is, however, of an entirely empirical character. If a be the quantity of sorbate per g of sorbent, then the FREUNDLICH isotherm is expressed by the equations

$$a = \alpha p^{1/n} \text{ or } \log a = n^{-1} \log p + \log \alpha \qquad (9)$$

α is a constant representing the quantity of sorbate at $p = 1$ and n is another constant characteristic for the system investigated. According to (9) $\log a$ is a linear function of $\log p$. Plotting observations of $\log a$ against $\log p/a$ straight line should arise and the value of $1/n$ may be read from its slope.

A convenient manner to see how $1/n$ affects the slope of the isotherm, is to assume a saturation point where the quantity of sorbate is a_s at a pressure p_o:

$$\frac{a}{a_s} = \alpha \left(\frac{p}{p_o} \right)^{\frac{1}{n}} \qquad (9a)$$
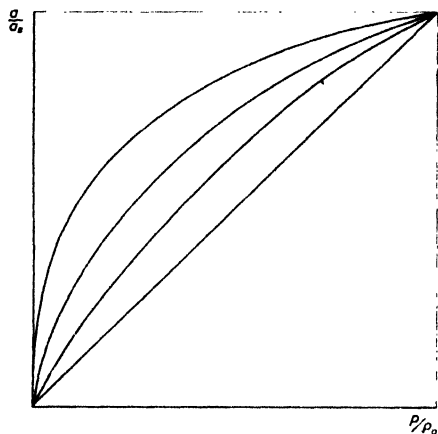
Fig. 20. FREUNDLICH isotherms.

The curves represented by this equation are schematically shown in Fig. 20. The straight line stands for $n = 1$, the other curves, all curves concave to the pressure axis, for increasing positive values of n. Values of $n < 1$ would give rise to curves which are convex to the p-axis.

Over a not too extended range of pressure or concentration the FREUNDLICH isotherm fits the observations reasonably well in a great many cases. This is however quite a common result, since equations of the form (9) can be made to fit almost any curve by selecting the proper value of n [2].

a. 3. *Some particular instances of sorption*

We shall forbear from giving a survey of the very extended literature on sorption phenomena and shall confine ourselves to treat in more detail a few characteristic examples. To that end four different types of well investigated sorbing systems will be selected, charcoal, silica gel, cellulose and nitrocellulose.

Amorphous carbon, though not being a gel in the usual sense, represents a characteristic sorbent. Its behaviour may serve to illustrate a number of characteric features of sorption. It is further thereby characterized that it does not show the phenomenon of swelling and that, according to all expectations, sorption will be exclusively or mainly a process of surface action.

Silica gel is a classical example of an irreversible "xerogel" it being formed by

---

[1] E. Hückel, *Adsorption und Kapillar Kondensation*, Leipzig 1928, p. 288.

[2] Cf. M. Reiner, *Gebrauch der Potenzfunktionen zur Darstellung einer naturgesetzlichen Beziehung*, *Naturwiss.*, 21 (1933) 294.

a process of drying from a true gel originating from a solution. Sorption on silica gel doubtlessly represents a case of 3-dimensional sorption, the sorbate penetrating throughout the mass of its body, though not giving rise to any appreciable amount of swelling.

Cellulose and particularly nitrocellulose represent reversible xerogels showing the phenomena of limited and unlimited swelling respectively. In certain instances sorption of a vapour may give rise to continuous transitions between the xerogel state and the formation of a true solution.

A thorough study of sorption in these characteristic instances may serve to reach a general understanding of sorption phenomena in other colloid systems too.

a Amorphous carbon. In Fig. 21 the sorption isotherms of some permanent gases in charcoal at low temperatures and pressures according to measurements of
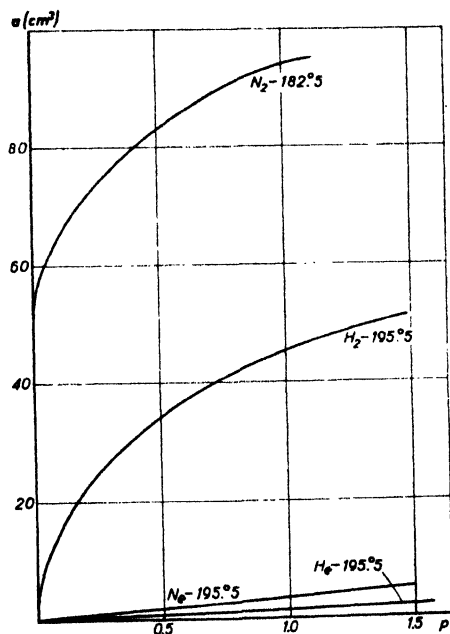
Fig. 21. Sorption isotherms of permanent gases on carbon. $a =$ cm³ gas ($0°$, 76 cm) absorbed; $p =$ pressure in mm.
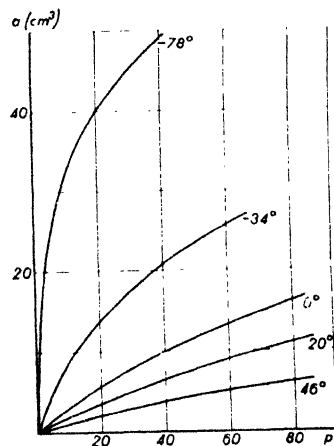
Fig. 22. Sorption isotherms of CO on carbon at various temperatures.

CLAUDE[1] are reproduced. These curves may obviously be interpreted as being of the LANGMUIR type (cf. Fig. 15). In conformity with the doubtlessly small heat of sorption of the inert gases, He and Ne, their isotherms represent straight lines of small slope.

It will be clear, that the shape of the isotherms will also depend upon the temperature, in particular upon the "reduced" temperature (fraction of the critical temperature). From the positive heat effect it follows that the sorption will decrease with increasing temperature. This is illustrated by Fig. 22, showing

---

[1] CLAUDE, *Compt. rend.*, 158 (1914) 861.

the isotherms of the system carbon monoxide—carbon at rising temperatures[1].

A logarithmic plot of these observations according to equation (9) yields straight lines in all cases.

All gases give straight isotherms analogous to HENRY's law if a sufficiently high temperature is selected e. g., hydrogen and nitrogen above — 80°; CO_2 and CH_4 above 100°, even water vapour above 220° (at pressures below 700 mm)[2].

According to well known principles, a process due to a positive heat effect is counterbalanced by a rise in temperature and in these cases, where only "secondary valence forces" play a part, the heat effect itself will be also smaller at higher temperatures.

Mention was already made in a former section that also the sorption of SO_2 on amorphous carbon at — 10° is very well represented by a LANGMUIR isotherm up to a relative vapour pressure of about 50 % where more than 0.9 g of SO_2 is taken up by 1 g of carbon[3]. The complete isotherm of this system is shown in Fig. 23a; Fig. 23b shows a plot of p/a against p which, on account of equation (2) (page 516) should represent the linear equation:

$$\frac{p}{a} = \frac{1}{a_s b} + \frac{p}{a_s}$$
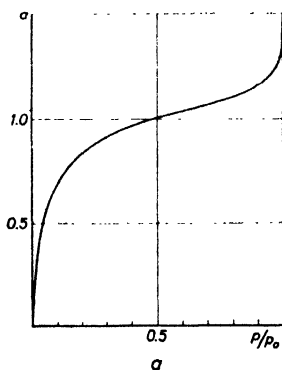
Fig. 23a. Sorption isotherm of SO_2 on carbon (— 10°)
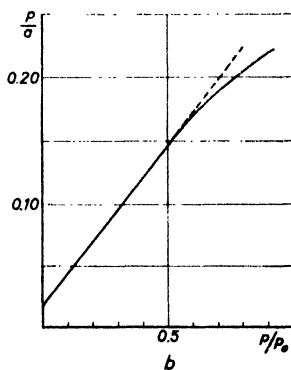
Fig. 23b. Observations from Fig. 13a (plot of p/a against p/p_o)

Though the conditions of the experiment doubtlessly fall beyond those theoretically presumed in deriving the LANGMUIR formula (monomolecular layer, ideal gas), a considerable part of the isotherm conforms with this equation.

According to section α (p. 514) this means that the affinity of the process remains sensib: constant no matter if mono or multilayers are adsorbed on the surface of the sorbent. Apparently the

[1] MISS HOMFRAY, Z. physik. Chem., 74 (1910) 129, 687.
[2] These data refer to purified ash free carbon.
[3] A. MATTHES, Kolloid-Z., 108 (1944) 81 according to observations of A. M. WILLIAMS, Proc. Roy. Soc. Edinburgh, 37 (1916/17) 161.

drop of the binding energy from the first to the next layers of adsorbed molecules is just compensated by an increase of the entropy of the sorbate, which is — qualitatively at least — a very acceptable assumption.

At relative pressures above about 0.8, the isotherm bends up and becomes convex to the pressure axis. This reminds us of Fig. 18 (p. 521) and should be interpretated as capillary condensation, starting toward attainment of the saturation pressure.

Let us now see, if the energetic effect conforms with this picture of the process. In Fig. 24 the differential heat of sorption *q-diff* as a function of the quantity of sorbate is shown[1]. (This is the heat given out if one g of SO_2 is absorbed by an infinite amount of the sorbent). On the ordinate the ratio between *q-diff* and the latent heat of evaporation of SO_2 (93 cal/g) is recorded.

It will be seen that the heat effect remains positive throughout the entire process, but becomes very small beyond about $a \infty 1$. Comparing this with Fig. 23a and 23b, it appears that this corresponds to the value of a, where the isotherm begins to deviate from the LANGMUIR type and where it has its inflexion point. From thereon, the smaller heat effect, due to capillary condensation, begins. The dotted lines in Fig. 24 correspond to the heat effect calculated by HÜCKEL for such a process on behalf of various assumptions.

Quite a similar behaviour is found if the sorption of the vapours of organic liquids like alcohol, benzene and carbon bisulphide by carbon is investigated[2].

The influence of temperature on the isotherms is in conformity with the picture: higher temperature reduces the quantity of sorbate in the lower pressure range, but has no influence on the quantity sorbed at the saturation pressure of the liquid. This was shown by HÜCKEL from observations of BROWN[3]. The last mentioned fact is exactly according to expectations if a process of capillary condensation is involved, since the capillaries will be filled at saturation. Furthermore, the calculated liquid volume of the sorbate at saturation is almost independent of the kind of substance, as is shown by the following table, selected from measurements of GURWITSCH[4].
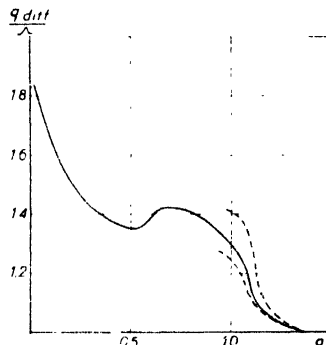
Fig. 24. Differential heat of sorption as a function of the quantity of Sorbate.

TABLE 3

WEIGHT AND CALCULATED LIQUID VOLUME OF VARIOUS SUBSTANCES ABSORBED BY ANIMAL CHARCOAL FROM SATURATED VAPOUR AT 18°.

Substance	g	$v = g/\varrho_t$
ethyl acetate	0.309	0.335
chloroform	0.500	0.327
benzene	0.309	0.351
carbon bisulphide	0.475	0.366
water	0.328	0.329

[1] Observations of WILLIAMS, *loc. cit.* (Figure borrowed from the book of HÜCKEL, p. 274).
[2] Cf. A. D. LAMB and A. S. COOLIDGE, *J. Am. Chem. Soc.*, 42 (1926) 1146.
[3] B. E. BROWN, *Phys. Rev.* (2), 17 (1921) 700; cf. E. HÜCKEL's book, p. 284.
[4] L. GURWITSCH, *J. russ. phys. chem. Ges.*, 47 (1915) 805.

Though the quantities absorbed at lower equal values of relative vapour pressure are very markedly different, they are almost equal at saturation pressure. Even water, which, in other respect, shows a very different behaviour towards amorphous carbon, as will be discussed now, makes no exception.

Very good information on the sorption of water vapour by amorphous carbon may be derived from a paper of COOLIDGE[1]. This author showed that the amount of water taken up at relative vapour pressure below about 0.3 is very small, if the carbon is cautiously freed from inorganic impurities which tend to increase water absorption very considerably in that range. In the experiments further referred to here, pure sugar charcoal was used. The isotherms measured at 0° and 100° are reproduced in Fig. 25.

Fig. 25. Sorption isotherms of water vapour on sugar charcoal.

It is seen, that the curves are entirely different from those of SO_2 (Fig. 23a). We shall first discuss the initial part below a rel. spresure of 0.3, where the quantity of water absorbed remains below 10 mg per g of carbon. This affords a striking contrast with the sorption of a substance like benzene of which almost 50% of the saturation quantity is already absorbed at a rel. pressure of 0.001. The "affinity" of carbon as a "hydrophobic" substance is small towards water and considerable towards many organic compounds.

COOLIDGE very carefully investigated the sorption in the low pressure range below $p/p_0 = 0.1$ and showed that the isotherms here represent practically straight lines. According to the preceding sections it is thereby demonstrated that the heat of sorption cannot have an appreciably positive value. COOLIDGE proved that the temperature coefficient of sorption is negative in this region, as shown in Fig. 26, referring to the low pressure region exclusively. At equal values of p/p_0 sorption is favoured by increasing the temperature, a very unusual phenomenon indeed. This proves that the process is endothermic and must be explained by high entropy of the absorbed water. A recent paper of LOISY[2] affords support to this point of view, showing, by measurements of specific volume, that a very small quantity of the absorbed water must be intimately mixed with the carbon.

At higher relative pressures, the isotherms in Fig. 25 become convex against the pressure axis and the 0° isotherm crosses the 100°-one, the temperature coefficient simultaneously changing its sign. The heat of sorption now becomes positive but remains very small, though increasing somewhat with rising pressure. At higher temperatures, the heat effect is smaller than at lower temperatures. COOLIDGE tends to explain the rapid rise of the isotherms in the range of intermediate pressures by the formation of multilayers on the carbon surface due to the polarity and the association
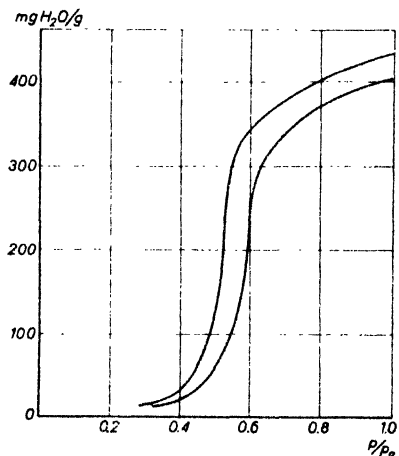
[1] A. S. COLLIDGE, *J. Am. Chem. Soc.*, **46** (1924) 596.
[2] M. R. LOISY, *Bull. Soc. Chim.*, (5) **8** (1941) 654.

of the water molecules (cf. also page 521). Capillary condensation which — for water — comes into consideration beyond $p/p_o = 0.3$ (see page 522) may, obviously, be equally well presumed. The second point of inflexion of the isotherms, followed by their final part, which is now concave to the pressure axis, however, then requires a special explanation. As such, a similar reasoning may be considered as will be set forth later for the corresponding feature of the system silica gel-water.

β *Silica gel*. For sake of comparison, we shall begin to show a diagram borrowed from COOLIDGE, representing the sorption isotherms of water on two specimen of carbon, silica gel, a platinum surface and a glass surface (Fig. 27). On the ordinate the quantity of sorbate, expressed in per cent of the saturation value, is recorded; the abcissa stands for the relative vapour pressure. For platinum and glass, the isotherms are characteristic for surface condensation. The isotherm of silica gel shows a new characteristic feature, viz., a hysteresis loop. Upon sorption and upon desorption the curves take a different course in a certain range of pressures. Next, in Fig. 28, the sorption isotherms of water, ethanol and benzene are shown[1]. The quantity of sorbate is here expressed in ccm per gram silica gel for reasons which will become clear below. On the abcissa the absolute vapour pressure in mm Hg is recorded. The endpoints of the curves O_3 correspond to saturation pressure. The three curves are strikingly similar, only showing characteristic differences in the low pressure range below point O_1.
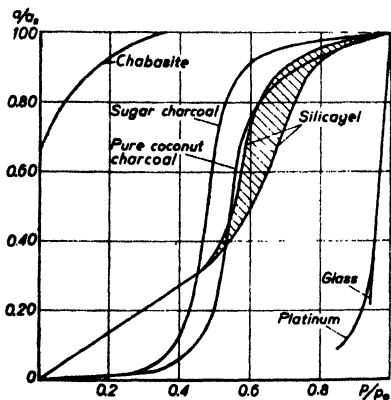
Dry silica gel is a hard transparent and apparently amorphous material[2]. No observable change in volume takes place upon

Fig. 26. Sorption isotherms of water vapour on sugar coal in the low pressure region at various temperatures.

Fig. 27. "Reduced" isotherms of water on various sorbing substances; sugar charcoal, pure coconut charcoal, silica gel, glass, platinum, chabasite.

sorption of either water or other substances. Other interesting phenomena are,

---

[1] J. S. ANDERSON, *Z. physik. Chem.*, 88 (1914) 91; *Thesis*, Göttingen 1914 (Fig. 18, borrowed from E. HÜCKEL).
[2] It is prepared by drying the coherent hydrogels obtained upon acidification of sodium silicate solutions under suitable conditions.

however, observed[1]. If silica gel, after being dried in vacuo over sulphuric acid at room temperature[2], is gradually loaded with any of the substances referred to in Fig. 28, the outward appearance of the gel remains unchanged untill point O_1 is reached. The sorption isotherms are concave to the pressure axis (in the case of water this part of the curve is almost straight — cf. Fig. 27 — which is not well represented in Fig. 28). Upon further increasing the vapour pressure, the isotherm bends off and becomes concave to the p-axis until about point O_2, beyond which higher pressures have very little effect on the quantity of sorbate[3]. If the vapour is now pumped off and the pressure lowered, the isotherms take a somewhat other course between O_2 and O_1, showing a sudden change of direction in O (compare also Fig. 31). The sections O_0O_1 and O_2O_3 are, however, reversible, i. e., they are identical at sorption and desorption.

Between O_1 and O_2 a peculiar optical effect is observed, the gel, which is perfectly transparent along O_0O_1 and O_2O_3 becomes opaque or turbid in the region of the hysteresis loop between O_1 and O_2.

A further characteristic feature is that, as shown in Fig. 28, the points O_1, O_2 and O_3 correspond to almost equal volumes of liquid absorbed.

Fig. 28.   Sorption of water, alcohol and benzene on silica gel.

The explanation of the phenomenon of hysteresis and the optical phenomena, as given by ZSIGMONDY, is as follows. The gel should be visualized as a system containing a great many small capillary pores. Along part O_0O_1 (compare Fig. 28 and preferably Fig. 31) sorption is due to adsorption on the inner surface of the capillaries (or, eventually, also to some kind of solid solution in the substance of the gel). At higher pressure, beyond O_1, condensation of liquid in the capillaries begins (first, of course, in those with the smallest diameter). Arriving at point O_2 the capillaries are filled. The final section O_2O_3, where the quantity absorbed shows hardly any further rise at increasing pressure, corresponds, according to ZSIGMONDY, to the final filling of the menisci of the capillaries which end on the surface of the gel. ZSIGMONDY also assumed that the walls of the capillaries are uncompletely or unevenly

[1] The subject was first studied in the classical work of VAN BEMMELEN (loc. cit., p. 534), later by ZSIGMONDY et al., Z. anorg. Chem., 71 (1911) 356; 75 (1912) 189 and many others.

[2] The gel then still contains about 5% water, which can only be disengaged by heating at very high temperatures and may be regarded as chemically bound water of constitution.

[3] Beyond O_3 arbitrary quantities of liquid can, of course, be condensed on the gel which would imply a further vertical rise of the curves which is omitted here.

wetted (cf. Fig. 19b, p. 523) during the filling process. The vapour pressure is, therefore, greater than complete wetting would imply.

On reversal of the process, now lowering the vapour pressure, curve O_3O is followed. At O, evaporation from the hollow menisci of the capillaries sets in and the curve rather abruptly changes its direction, the pressure, for a while, remaining almost constant at the value corresponding to the radius of the capillaries, according to equation (8), (p. 522). The capillary walls, now being completely wetted, the pressure at which the capillaries are emptied is lower than that observed during filling. In this way the hysteresis loop was explained.

The fact that the liquid volume of the sorbate, corresponding to the points of inflexion O_1 and O_2, was found to be almost exatly equal for different substances sorbed on the same sample of silica gel[1], afforded strong support to this theory. So did the discovery of ANDERSON, that the radius r of the capillaries, calculated from the pressure in point O according to equation (8), yielded equal values for different substances. (13 — 28 Å in various samples of silicagel)[2].

The optical phenomena (opaqueness of the gel in the region of the hysteresis loop) may be explained thus: the pores are so minute as compared to the wave length of light that they do not give rise to scattering. The dry as well as the completely wetted gel remain, therefore, completely transparent. In the hysteresis area, however, where the pores are partially filled with liquid and partially with vapour (or air), larger regions of the gel containing emptied pores exist next to larger regions, where the capillaries are still filled with liquid[3]. This now gives rise to optical discontinuities whose dimensions may fall within the order of magnitude of the wave-length of light and can be set responsible for the opaqueness of the gel.

In later years various observations were published throwing doubt as to the correctness of ZSIGMONDY's theory. Numerous sorption isotherms on silica gel were found not showing any hysteresis at all. Sometimes water yielded isotherms of the type discussed hereabove and other substances, sorbed on the same gel sample, did not[4].

According to Mc GAVACK and PATRICK[5] the hysteresis completely disappears if all traces of air are removed from the gel. PATRICK still stuck to the theory of capillary condensation, but failed to explain how the presence of traces of permanent gases could exert such a large influence on the character of the sorption isotherms. Later he reported that air does not affect the shape of the isotherms.

The situation seems to be very satisfactorily elucidated by the recent investigations of COELINGH[6]. She systematically studied the interference colours in the thin surface layers on old weathered glass walls and the changes occurring upon sorption and desorption of various vapours. She proved conclusively that capillary condensation is involved in these phenomena and could perform equilibrium measurements. On basis of her findings she added new ideas to the theory of capillary condensation, which now seem to clear up the remaining difficulties.

---

[1] BACHMANN, Z. anorg. Chem., 79 (1912) 202.
[2] ANDERSON, loc. cit., p. 529.
[3] This will be better understood later (See Fig. 30).
[4] LAMBERT et al, Proc. Roy. Soc. London, 117 (1928) 183; 122 (1929) 497; 134 (1932) 246; 136 (1932) 363; 144 (1934) 205; 153 (1936) 584.
[5] J. Mc GAVACK and W. A. PATRICK, J. Am. Chem. Soc., 42 (1920) 946; compare also W. A. PATRICK, Kolloid-Z., 36 (1925) 272. (ZSIGMONDY Festschrift).
[6] M. B. COELINGH, Thesis, Utrecht 1938; Kolloid-Z., 87 (1939) 251; cf. also L. H. COHAN, J. Amer. Chem. Soc., 60 (1938) 433.

According to COELINGH, the different results obtained by various authors as to the shape of isotherms of silica gel depend on the nature and the previous history of the samples used. "Activated samples" (exposed to high temperatures for some time to remove last traces of water), as they were used by ANDERSON and PATRICK, contain smaller pores, than the freshly prepared not "activated" samples employed by ZSIGMONDY and coworkers. She gave the following survey, including also two silicic minerals: hydrophane, containing very large pores and chabasite (belonging to the class of zeolites), containing extremely small pores with a diameter of 3.6 Å, as known from X-ray analysis.

TABLE 4

SPECIMEN OF SILICA AND THEIR AVERAGE PORE RADII IN $m\mu$ ACCORDING TO COELINGH

Group 1	Group 2	Group 3	Group 4
Hydrophane; Aged silica gel	Freshly prepared silica gel	activated silica gel	Chabasite
V. BEMM. $\Big\}\ r > 40$ BACHMANN	ANDERSON $\quad r = 1.3—2.8$ COELINGH $\quad r = 6.0—13.4$	LAMBERT $\quad r = 0.8—1.0$ PATRICK $\quad r = 0.1—0.6$	$r = 0.18$
no inflexion points no hysteresis with various substances	inflexion points hysteresis with various substances	inflexion points and hysteresis with *water only*	no inflexion points no hysteresis

The pore radii in the groups 2 and 3 were determined from the vapour pressure in point O of the isotherms, those of group 1 could be estimated from observations of VAN BEMMELEN. Specimens having a pore radius between 1.5 and 15 $m\mu$ (group 2) yield the typical isotherms exhibiting hysteresis and inflexion points with all volatile substances investigated. Those with a pore radius of about 1 $m\mu$ behave towards water as group 2 and towards other liquids as group 4. The isotherm of chabasite and water vapour (see Fig. 27, curve 6) is of the LANGMUIR type and reveals a great affinity between sorbent and sorbate. Absorption of water in pores of so small dimensions can, of course, not be termed capillary condensation; it rather represents a typical case of "solid solution". Other substances as e.g., benzene are not absorbed by chabasite, since the pores are too small to allow of permeation of these larger molecules. Gel specimen with pores larger than 40 $m\mu$ (group 1) do not give rise to hysteresis either and yield isotherms convex to the pressure axis without hysteresis. Obviously, the occurrence of inflexion points is always connected with the phenomenon of hysteresis[1].

It is clear, says COELINGH, that capillaries of large diameter fail to show the typical phenomena of capillary condensation, since condensation occurs at pressures lying so near the saturation pressure that they cannot be measured. Very small capillaries, with a diameter falling within the range of molecular dimensions, can but

---

[1] Hysteresis and inflexion points have also been observed in certain specimen of amorphous carbon, when sorbing water. The pore dimensions of the latter, accordingly, fall into group 3 as COELINGH showed.

show adsorption. Only capillaries of the size belonging to group 2 will be capable to give rise to the typical phenomena of capillary condensation with hysteresis. The lower limit of r for water corresponds to a somewhat smaller value than that for other substances like benzene and alcohol[1].

Another valuable contribution of COELINGH is her explanation of hysteresis, which is far more plausible than ZSIGMONDY's assumption of incomplete and complete wetting of the capillary walls. From her own experiments she could deduce that the capillary radius, calculated by using equation (8) (p. 522), was always exactly twice as high if derived from the *ab*sorption curve then if derived from the *de*sorption one. This led her to the conception already illustrated in Fig. 19a and 19d (p. 523). Since this simple picture would, however, imply a distinct inflexion point on the absorption curve too, a modification was introduced explaining the typical differences observed between the absorption and desorption curves.

Fig. 29. Condensation in a capillary of oval cross section after COELINGH.

COELINGH points out that the cross-section of the pores in an actual gel will hardly ever be circular. She discusses what would happen if the cross-section were oval (see Fig. 29).

Obviously, condensation will begin where the curvature of the wall is highest. With increasing pressure, condensation proceeds until the water layer has attained a cylindrical surface. A further rise of the pressure will now rapidly fill the whole capillary (since further growth of the water layer would now *lower* the vapour tension instead of increasing it).

Now, assuming a spreading of the degree of excentricity of the pores in the gel. the absorption curve can be explained thus (compare Fig. 31):

Fig. 30. Scheme explaining turbidity in the hysteresis area.

O_0O_1 the pores are empty but for the adsorption layer on the walls;

O_1 in the pores with largest excentricity condensation begins;

O_1O_2 the pores are gradually filled, the vapour ensiont rising gradually according as the curvature of the liquid surfaces in the pores (Fig. 29) becomes smaller;

O_3 all pores are filled. The pressure in O_3 corresponds to the curvature of the cylindrical surface shown in Fig. 29.

Along the desorption curve the emptying of the pores begins at O, corresponding to the pressure of the half spherical menisci. The pressure remains constant until emptying has so far proceeded that the curvature of the layers remaining on the capillary wall has become higher than corresponds to that of the spherical menisci in O.

If we, finally, realize that the pores in the gel will certainly communicate and will represent anastomising spaces of irregular shape rather than a series of tiny

---

[1] This fact was also explained by COELINGH, but it would lead us too far afield to go into this matter here.

mutually independent little tubes[1], we can also understand the optical phenomena observed in the hysteresis area. In Fig. 30 a schematic picture of the gel frame-work is given. When evaporation from the filled gel begins, a considerable negative pressure, due to the capillary pull of the hollow menisci, will be generated in the liquid of the gel[2]. In the interior of the gel, here and there occasional "breaks" will occur, particularly, if dissolved gases are present which may give rise to little gas bubbles acting as "germs". Then regions far larger than the size of the simple pores and containing vapour (see areas surrounded by dotted lines in Fig. 30) will alternate with others still filled with liquid.

BÜTSCHLI[3], experimenting with alcohol -gels of gelatin, which also show simliar optical phenomena upon drying, has observed, using large blocks of gelatin gel, that the turbidity first appears in the centre of the blocks. The turbid area then spreads out according as the alcohol is evaporated. The outer skin of the blocks remains transparent. We may presume that the outer parts of the objects, have the tendency to shrink a little more than the inside parts and hence adopt a denser structure. "Breaks" of the liquid, apparently, preferably occur in the least compressed regions.

BÜTSCHI observed air bubbles being expelled, if opaque blocks were laid in water, thus actually demonstrating the presense of air. A similar observation with opaque silica gel was reported by VAN BEMMELEN[4].

It becomes necessary now, to discuss another item, viz., the points of view which may be derived from studying the preparation of the material called silica gel. To that end we may base ourselves upon the classical investigations of VAN BEMMELEN.

Fig. 31.   Vapour pressure curves according to VAN BEMMELEN.

Silica gel is prepared by acidifying aqueous solutions of sodium silicate under suitable conditions. The freshly prepared "primary" gels formed by the gelation of the solution have, of course, a very high water content, depending on the concentration of the solution. The gel is then washed and dried. Curve a in Fig. 31 shows how the vapour pressure of the freshly prepared gel changes during drying and

---

[1] Compare also the photograph of the V_2O_5 gel reproduced in Fig. 3b, p. 493, it being understood that the average pore diameter in dry silica gel is less than 0.1 of that of the largest pores shown there.

[2] It should be recalled that in a capillary with a radius of 2.5 mμ water reaches a height of 10000 m.

[3] O. BÜTSCHLI, Untersuchungen über Strukturen, Leipzig 1898; Uber den Bau quellbarer Körper, Göttingen 1896.

[4] J. M. VAN BEMMELEN, Die Absorption, p. 327.

that the curve joins the known vapour pressure diagram of the dry gel at O. The curve a represents an irreversible branch of the diagram. It cannot be followed in opposite direction. Along this curve the gel shows considerable shrinking, its volume decreasing according as water is removed. (Upon adding water no swelling is observed). The voluminous open network structure present in the original gel contracts and "collapses" during the process of shrinkage (Cf. page 571). At point O this process comes to a sudden end; the gel framework has reached a state of compression not allowing further contraction[1]. From this point onward the volume remains constant and the vapour pressure curve shows a sudden change of direction. Upon further extraction of water the desorption curve of capillary condensation (OO_1) begins. It will be clear that the position of point O must depend on the constitution of the fresh primary gel. VAN BEMMELEN showed that it changes position according as the conditions of preparation of the latter were varied. Very voluminous gels, prepared from dilute sodium silicate solutions, could contract further, and point O was reached at a lower water content, than if the gel was prepared from more concentrated solutions. Gels, aged before the drying process began, reached O at a higher water content than freshly prepared ones. This is in conformity with the general expectation that diluted and freshly prepared gels will have a finer, more loosely built and more flexible network frame than concentrated and aged ones.

The water content on the branch O_0O_1 of the diagram is the higher according as the original dilution of the gel was greater, indicating that more and finer capillaries are present. Upon concentration and ageing, the framework, apparently, becomes coarser.

A diagram, based on VAN BEMMELEN's observations, showing the change of the isotherms according as the age of the gel increases, is reproduced in Fig. 32. It demonstrates that the total pore volume and the coarseness of the pores increases with age. For sake of comparison the isotherm of the mineral hydrophane has been added. It is seen that the behaviour of the oldest gel approaches that of the mineral substance; hysteresis disappearing in both cases.

Fig. 32. Diagram showing the change of the isotherms as a result of ageing.

In other gels, such as gelatin, cellulose and agar-agar, the gel frame seems to be more flexible and shrinking does not stop before all the water is extracted from the gel, though, as could be demonstrated in the case of cellulose and gelatin (cf. page 573), other phenomena indicate that still a similar effect exists to a less marked degree. It can, however, be enlarged by certain artificial means even to such an

---

[1] According to VAN BEMMELEN the end of shrinking does not exactly coincide with point O, but comes a little later.

extent that the lack of shrinking power combined with the phenomenon of turbidity becomes very apparent (see p. 582).

Finally, we must briefly consider the question what may be the mechanism of shrinkage during the extraction of water from the primary gel. This apparently self evident phenomenon is actually not fully clear and still open to discussion. There, is one theory stating that the capillary pull exerted by the menisci of the pores ending on the surface of the gel, giving rise to, so to speak, a negative tension throughout the body of the gel, is mechanically equivalent with a 2 dimensional compression tendency of the gel surface. These facts would be responsible for the shrinkage proceeding according as water is removed by evaporation, until the framework is so far compressed that it resist further contraction. Recalling that a certain amount of shrinkage of many primary gels can also be reached by transferring the aqueous gel to alcohol, one should be aware that other factors may also interfere and the question can, as yet, not be regarded as being satisfactorily settled.

The behaviour of silica gel upon extraction of water from the freshly prepared diluted gel can be visualized as a typical example of the shrinkage of an irreversíbel gel with a rather stiff framework, which — at least — remains rigid after the critical point O has been reached, resisting further volume changes of any appreciable extent.
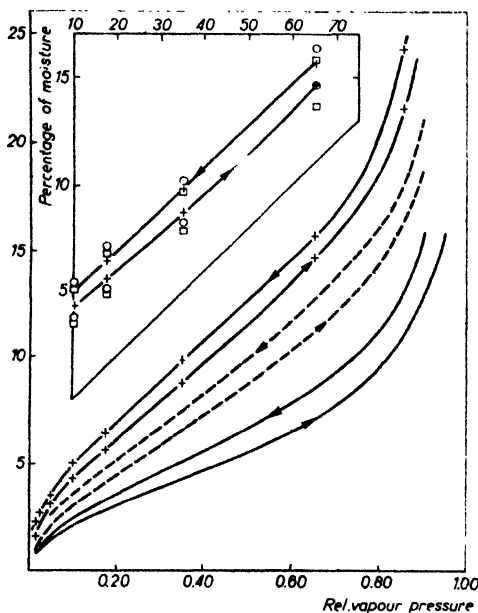
Fig. 33. Absorption and desorption isotherms of the system cellulose water. ———— native cotton; - - - - - - mercerised cotton (25°); + isotropic filaments. In a separate graph part of the isotherms of two specimens of viscose rayon (o and □) and isotropic filaments (+) are given on another scale.

γ Cellulose. We shall confine our selves to the important system cellulose-water, since the absorption of other substances than water is either very small or not well investigated[1].

In liquid water, cellulose shows limited swelling. The amount of water taken up depends on the kind of sample. Native cellulose, such as purified cotton-, and ramie fibres, take up about 0.5 g water per g of the dry material, artificial fibres 1.0 g or even more. The sorption isotherms of various samples show a similar difference; at a given vapour pressure native fibres absorb less water than regenerated ones[2]. The shape of the isotherms is, however, the same in all cases. A very characteristic feature is,

---

[1] Ammonia is a substance which is also readily absorbed at temperatures below zero.

[2] The vapour pressure isotherms of cellulose were extensively studied a. o. by A. R. URQUHART and A. M. WILLIAMS, *J. Textile Inst.* 15 (1924) T 138; 16 (1925) T 135; A. R. URQUHART and N. ECKERSALL, *ibid.*, 21 (1930) T 499; J. G. WIEGERINK, *Textile Research* 10 (1940) 357; cf. also P. H. HERMANS, *Contribution to the Physics of Cellulose Fibres*, Amsterdam-New York 1946.

that different curves are obtained upon absorption and upon desorption. There is a marked hysteresis extending itself over the entire range of relative pressures.

In fig. 33 *isotherms* of various specimens at 20° are reproduced.

At low vapour pressures, the curves are concave to the pressure axis, then remain almost straight for a while and, at higher vapour pressure, become convex to the pressure axis. For all artificial fibres, (consisting of cellulose precipitated from a solution), including isotropic model filaments prepared from viscose[1], the curves are almost identical.

URQUHART and WILLIAMS (*loc. cit.*) showed that the ratio of the water content of different samples at a given value of p/p_o remains practically constant over the entire range of pressures. Taking the isotherms of cotton as a standard, the isotherms of other samples may be characterized by this ratio as a parameter, termed "sorption ratio". This simple relation holds, however, only in the range between 5 and 65% relative humidity. Beyond this range deviations were observed, which will be discussed later. The magnitude of hysteresis, which can be expressed by the quotient of the absorption and the desorption ratio, does not show great variations from one fibre specimen to another. A survey is given in Table 5.

TABLE 5

SURVEY OF ABSORPTION AND DESORPTION RATIOS (TAKEN BETWEEN 5 AND 65% OF REL. HUMIDITY)

	Abs. ratio	Des. ratio	Average	$\dfrac{\text{Abs. ratio}}{\text{Des. ratio}}$
Mercerised cotton	1.46	1.50	1.48	1.03
Mercerised ramie	1.53	1.64	1.59	1.07
Isotropic model filaments	2.02	1.98	2.00	0.99
Orientated model filaments . . .	1.94	1.90	1.92	0.98
Viscose rayon I	1.84	2.13	1.99	1.16
Viscose rayon II	1.79	2.03	1.91	1.13
Lilienfeld rayon I	1.95	2.08	2.02	1.07
Lilienfeld rayon II	1.89	2.01	1.95	1.06
Average of artif. filaments . . .	1.91	2.02	1.97	1.05

(The rayon samples indicated I and II differ from each other by the degree of their orientation; II being better orientated than I). Mercerised native fibres (fibres previously treated with sodium hydroxide) take an intermediate position between native cotton and regenerated fibres; so does woodpulp (not included in the table)

At first, most authors tended to explain the sorption process by assuming an adsorption on the "inner surface" of the fibres, followed by a capillary condensation, which is, indeed, in conformity with the general shape of the sorption isotherms. This interpretation would, hence, imply practically the same elements as those previously met with in the system silica gel water. It is to be noted, however, that there are some fundamental differences.

1. Whereas the volume of silica gel remains constant from 0 to 100% relative humidity, the volume of cellulose varies over this entire range according as its moisture content changes.
2. The density of silica gel differs considerably from that of crystalline SiO_2, thus allowing for a considerable pore volume of at least 60%. This is also in conformity with the fact that silica gel absorbs equal volumes of almost any liquid. The density of native and regenerated cellulose

---

[1] Model filaments are coarse horsehair-like fibres.

differs only very little from that of crystalline cellulose (2 and 4½% respectively) and cellulose is very selective towards other liquids, most organic liquids being not absorbed at all.

Hence, if "capillary condensation" occurs in water vapour at high relative pressures, the capillaries must be formed during the sorption process itself, since there exists no corresponding pore volume in the dry material. Furthermore the phenomenon of hysteresis in cellulose is, obviously, of another nature than that in silica gel and will require another explanation.

At the present moment we have at our disposal satisfactorily founded information on the structure of cellulose and a detailed picture of fibre structure which would seem to correspond rather well with the truth. On account of this knowledge, which is much more advanced than that on the actual structure of silicic acid gels, we may endeavour to give another more adequate interpretation of the sorption phenomena in cellulose [1].

According to § 8 c. 1 δ (p. 611) dry cellulose may be regarded as consisting of an intimate mixture of a crystalline and a non crystalline (amorphous) component. The affinity of cellulose for water must, in some way, be connected with the numerous hydroxyl groups in the molecule (three per glucose residue). It is even remarkable that cellulose is not completely soluble in water. Other linear polymers bearing "hydrophilic" groups e. g., polyacrylic acid, polyvinylalcohol and some polymeric carbohydrates are actually soluble in water. The explanation of the problem must be sought in the crystal structure. The molecules in the lattice are so arranged that very strong lateral bonds (of the nature of so called "hydrogen bonds") are formed between the OH-groups of adjacent molecules [2]. The crystalline modification occurring in native cellulose is entirely indifferent towards water. Its X-ray diagram undergoes no change whatever upon moistening and a cellulose monocrystal would be completely insoluble and exhibit no swelling. In an aggregate of less well ordered chain molecules (amorphous cellulose) such a tight cross linking of the molecules cannot occur for sterical reasons, and water can get hold of the residual affinities of the hydroxyl groups, penetrating between the chains and causing swelling. Amorphous cellulose should in a sense be considered as being soluble in water. Common cellulosic objects are non soluble and show only limited swelling in water, because the amorphous and crystalline portions are intimately connected with each other and the latter form cross-links of a permanent water resistant character, thus inhibiting the former from drifting apart [3].

The importance of regular lattice formation in causing insolubility is strikingly illustrated by the fact that substitution of some of the hydroxyl groups by methoxy groups (partial methylation of cellulose) gives rise to water soluble methylcelluloses. The irregular distribution of some methoxy groups along the chains apparently inhibits the formation of a well ordered lattice and gives rise to a profound change of solubility conditions.

Now the existence of a sorption ratio which is constant throughout a large range of rel. vapour pressures, suggests that the nature of the structural component responsible for the sorption is the same in all preparations, its quantity, however, varying

---

[1] Compare the work of the author and his collaborators, loc. cit. (p. 539).

[2] Cf. K. H. Meyer and L. Misch, Helv. chim. acta, 22 (1939) 59; P. H. Hermans, J. de Booys and Chr. J. Maan, Kolloid-Z., 102 (1943) 169; E. Schiebold, Kolloid-Z., 108 (1944) 248.

[3] It is to be noted that dependence of solubility on the crystalline modification and on the degree of subdivision of the substance considered is not an exceptional phenomenon. A striking example is found in phosphorus pentoxide, known in two crystalline modifications, corresponding to a monomeric and a polymeric form. Only the former is strongly hygroscopic and readily soluble in water, particularly if the crystals are small. The latter is almost insoluble in water and even capable of gel formation when its lattice is distorted.

from specimen to specimen. There is a difference in degree but not in kind between various cellulose specimens. This can be readily understood, if we consider the process of sorption as one exclusively connected with the amorphous portion, adopting the point of view that its percentage varies from one fibre specimen to the other. As a matter of fact, this hypothesis has proved a very fertile one and many other physical properties of cellulose fibres can be satisfactorily interpreted on this basis.

It should, however, be kept in mind that no very precise limit can be traced between the crystalline and the amorphous parts of a fibre, since the picture developed intrinsically implies gradual transitions between the two. As a first approximation, we may, however, rely upon this dualistic subdivision and thus discriminate a swelling and a non swelling structural component, it being understood that these are tightly connected together, since individual chain molecules may extend through several adjacent crystalline and non crystalline regions.

The sorption isotherm will than be the sorption isotherm of amorphous cellulose, intimately connected with its swelling. The total amount of water bound at a given pressure and the total heat given out upon complete wetting (integral heat of sorption), will be proportional to the amount of amorphous fibre substance.

On the other hand, the differential heat of sorption of dry cellulose (i. e., the heat evolved if 1 g water is sorbed by an infinite quantity of cellulose) should be equal for all specimen. It is actually borne out by experiments that the integral heat of sorption and the sorption ratio of various specimen are nearly proportional to each other, but that the differential heat of sorption of native and regenerated cellulose is equal[1]. The experimental data further reveal that the percentage of crystalline matter is greatest in native fibres and markedly lower in regenerated ones. In the latter, it shows but very small variations and seems to be constant in first approximation.

Native fibres show an increase of sorptive power after mercerisation, this being due to an increase of the percentage of amorphous substance as a result of the treatment.

According to X-ray research, the modification of the crystallites in mercerised and regenerated fibres is different from that in native cellulose and is termed cellulose II. With the aid of a procedure discovered by KUBO (heating in polar liquids like glycerol and glycol), fibres consisting of cellulose II can be transformed into cellulose IV, a modification very much ressembling cellulose I[2]. Simultaneously, a certain amount of recrystallisation takes place, as may be deduced from the intensity of the X-ray diagram and the change in optical behaviour upon transformation[3]. Fibres thus treated actually show a considerable decrease of sorptive power and integral heat of sorption.

On the other hand, HESS and coworkers[4] have succeeded in destroying the crystalline order in cellulose fibres by mechanical treatment. (Complete disappearance of the characteristic X-ray interferences). As a result of the treatment, sorptive power and integral heat of sorption increase considerably[5].

Finally, before we revert to the sorption equilibria, mention must be made of the recent discovery that cellulose II forms *a true hydrate*[6]. If dry fibres, whose crystalline part consists of cellulose II, are exposed to air of increasing rel. humidity, the 101 interference in the X-ray diagram shifts to another position, corresponding to an increase of the lattice spacing perpendicular to 101 from 7.32 to 7.73 Å. This

[1] P. H. HERMANS, *Contribution to the Physics of Cellulose Fibres*, Amsterdam-New York 1946.
[2] Cf. K. HESS, H. KIESSIG, and J. GUNDERMANN, *Z. physik. Chem.*, B 49 (1941) 64.
[3] P. H. HERMANS, *loc. cit.*
[4] K. HESS, H. KIESSIG, and J. GUNDERMANN, *Z. physik. Chem.*, B 49 (1941) 64.
[5] P. H. HERMANS and A. WEIDINGER, *J. Am. Chem. Soc.*, 68 (1946) 2547.
[6] P. H. HERMANS and A. WEIDINGER, *J. Colloid. Sci.*, 1 (1946) 185.

shift is reversible and its magnitude depends unequivocally upon the vapour pressure, a phenomenon characteristic of macromolecular addition compounds (cf. p. 521). A discussion showed that a hydrate of low water content, corresponding to either $C_6H_{10}O_5 \cdot \frac{1}{3} H_2O$ or $C_6H_{10}O_5 \cdot \frac{1}{2} H_2O$, must be assumed.

This hydrate was termed hydrate I. Another crystalline hydrate of cellulose, which we shall designate as hydrate II and which, in contrast to the former, is never formed spontaneously when starting from dry crystalline cellulose, was discovered by two Japanese investigators[1]. It is obtained if other cellulose compounds containing water, such as sodium cellulose or cellulose xanthate, are decomposed at low temperatures. It·is characterized by a further shift of the 101 interference, corresponding to a spacing of 8.98 Å and was termed watercellulose by its discoverers. It probably contains 1 molecule water more than hydrate I and it is unstable at ordinary temperature. It decomposes spontaneously, loosing water, into hydrate I, a stable compound which is only further decomposed at very low vapour pressures.

The situation may be summarized thus: the lattice of cellulose I is not accessible to water. Here sorption of water and swelling is strictly confined to the amorphous portion of the fibre. Into the lattice of cellulose II a small amount of water can penetrate spontaneously and is stoichiometrically bound there by strong "chemical" forces. The free energy of hydration is, obviously, greater than the work necessary to widen the lattice. The remaining lattice energy is, however, great enough to prevent further water from penetrating and the free energy of hydration of the second hydrate is too small to overcome this cohesion. In cellulose II sorption of water vapour is, for a small part (corresponding to 0.03—0.05 g. water per g of cellulose), due to the formation of hydrate I. Further sorption is again a function of the amorphous portion.

Let us now see how the amorphous portion of the fibre is likely to take part in hydrate formation.

If the lattice cohesion of the crystallites has been overcome by some previous process, hydrate II can be temporarily formed as an unstable crystalline modification. Now the capacity of hydrate formation thus demonstrated should be considered as inherent to the cellulose molecule as such. A single chain molecule, if gradually exposed to increasing vapour pressures will first form hydrate I and then hydrate II[2]. So will a great deal of the chain molecules in the amorphous regions of cellulose fibres. Not being subject to order imposed by lattice formation, many chain sections may be considered as more or less "free". In contrast to the crystalline regions, the amorphous portion of the fibre will be capable of binding water in the form of hydrate II.

From the foregoing it must be concluded that in mercerised and regenerated cellulose the crystalline part will form hydrate I and the amorphous part both hydrate I

---

[1] I. SAKURADA and K. HUTINO, Kolloid-Z., 77 (1936) 347.

[2] The capacity for hydration of a substance in the crystalline state is a measure of its hydration capacity in solution. Sulphuric acid e. g., does form crystalline hydrates. In liquid mixtures of sulphuric acid and water, these hydrates, of course, are present as well. They are then dissolved in an excess of either water or sulphuric acid. In inorganic chemistry there are also numerous examples of solid substances forming chemically definite hydrates, though these cannot be detected by the common methods (X-ray diagram, isothermal or isobaric decomposition) which imply the existence of phases with crystalline order. According to R. FRICKE (WALDEN, Handbuch d. allgemeinen Chemie, 9 (1937), 520), a typical example is chromium (3) hydroxide. In such cases the vapour pressure isotherms do not represent the typical step-ladder type but resemble those of solutions (R. FRICKE, Naturwiss., 31 (1943) 469).

and II. In native fibres only the amorphous parts will form hydrates. In the lowest pressure range, where the formation of hydrate I is the dominating factor, this will give rise to a slight increase of the sorption ratio. This is exactly what was actually established [1]. At higher moisture contents, the small amount of water bound as hydrate I in the crystalline portion of cellulose II affects the sorption ratio to a negligible extent only. In this range the latter remains sensibly constant and proportional to the ratio of the amorphous fractions.

If to a dry cellulose fibre water is gradually added, hydrate formation will take place as a topochemical reaction. The water molecules will be bound at distinct places along the cellulose chains, mainly in the amorphous regions. This should give rise to a behaviour according to the LANGMUIR equation for the case of 3-dimensional sorption (p. 520). As is demonstrated by Fig. 33, this is in qualitative agreement with the experiments. In various ways it has been shown that the isotherms are also in quantitative conformity with the formation of hydrates.

Since two hydrates of different binding energy play a part, the isotherm should represent a superposition of two LANGMUIR isotherms with a different value of the affinity factor, a large one for the first hydrate and a small one for the second hydrate. From an investigation of MATTHES [2] it can be deduced that the sorption isotherm of regenerated cellulose can be quantitatively accounted for on this basis up to a relative humidity of abound 65%. MATTHES showed that $(p_1 - p_2)/(a_1 - a_2)$ should be linear in p (where p_1, a_1

Fig. 34. Verification of MATTHES' equation using the sorption data for cotton (absorption and desorption).

and p_2, a_2 are experimentally observed equilibrium values of vapour pressure and regain. Fig. 34 shows the verification of this relation.

The relative vapour pressure up to which MATTHES equation holds accurately, conforms to that where hydrate formation is completed. From density determinations and other facts it could be deduced that the amorphous fractions in native and regenerated fibers may be assumed to be about 0.4 and 0.75 respectively [3]. This leads to the following balance of water content if hydrate formation is carried to completion [4]:

Native cellulose.
water bound as hydrate   I in amorphous part .... $0.4 \times 3.7 = 1.5\%$
,,   ,,   ,, hydrate II ,,   ,,   ,, ... $0.4 \times 11.1 = 4.4\%$
                                        Total........................ 5.9%

Regenerated cellulose.
water bound as hydrate   I in whole fibre 3.7%
,,   ,,   ,, hydrate II in amorphous part .... $0.75 \times 11.1 = 8.4\%$
                                        Total........................ 12.1%

[1] P. H. HERMANS, loc. cit., (p. 539). A more detailed analysis of the subject is given there.

[2] A. MATTHES, Kolloid-Z., 108 (1944) 84. MATTHES was not aware of the formation of true hydrates and gave another interpretation of the sorption process which is, however, from a physical point of view not essentially different from ours.

[3] P. H. HERMANS, Contribution to the Physics of Cellulose Fibres, Amsterdam-New York 1946, p. 71; J. Polymer Sci., 1 (1946) 162.

[4] The composition of the hydrates is here assumed to be
$$C_6H_{10}O_5 \cdot \tfrac{1}{3} H_2O \quad \text{and} \quad C_6H_{10}O_5 \cdot 1\tfrac{1}{3} H_2O.$$

The totals of 6% and 12% respectively correspond to the water content reached by native and regenerated fibres at nearly 60% rel. humidity upon absorption.

The last part of the isotherms beyond 60—70% rel. humidity, which are convex to the p-axis and rise steeply towards saturation, were ascribed to capillary condensation by MATTHES and most earlier authors. Some serious objections against this interpretation have already been set forth on p. 538. To this may be added that all evidence points out that the water is homogeneously taken up between the chains in the amorphous regions to form a molecular dispersion. At 60% rel. humidity, the specific volume regenerated fibres is only about 20% higher than that of crystalline cellulose and the maximum diameter of the capillary spaces between the cellulose molecules can then be estimated to be in the order of 2 Å. It is, hence, physically senseless to assume capillary condensation. Even at saturation, the degree of swelling reached leaves hardly any room to speak of capillary condensation, apart from the fact that is it also very doubtful what capillary condensation would mean in a system of single chain molecules surrounded by water.

A theoretical interpretation of the behaviour of the system in the higher pressure range should, therefore, rather start from the theory of solution. Beyond 60% rel. humidity water dissolves in the amorphous regions. In this region there is either no thermal effect of sorption or a very small one and the free energy of swelling may entirely or mainly depend on entropy just as in ordinary solutions.

For the system wool-water a theory of this kind has been developed by CASSIE [1] and in a particularly interesting paper by HAILWOOD and HORROBIN [2]. The latter postulated formation of a monohydrate, i. e., binding of one water molecule by each monomeric residue of the fibre portion taking part in absorption. For the isotherm they gave the equation:

$$\frac{M}{m}\, a = \frac{a\beta h}{1 + a\beta h} + \frac{h}{1 - ah} \tag{10}$$

where M and m are the operative molecular weights of the monomeric residue and that of water and a is the total grams of water absorbed per gram of the dry polymer; h is the relative vapour pressure and a and β are constants. Comparing this with the experimental data the constants M, a and β can be computed.

It is seen from equation (10) that the quantity of water absorbed is considered as being given by the sum of two terms, one corresponding to a LANGMUIR isotherm (cf. equation (2) on p. 516) and another corresponding to RAOULT's equation for ideal solutions.

When applied to the systems wool-water and cellulose-water the values of M deduced from the experiment are larger than that of the actual monomeric residue (162 in the case of cellulose). In the sense of the theory this means that a certain fraction of the monomeric residues $(M - R)/M$ does not take part in absorption, where R is the actual molecular weight of the residue. In other words the percentage of crystalline substance can be deduced from the observations. This leads to the following figures: Wool 44—48%, natural silk 80%, native cotton 68%, regenerated cellulose 35%. The last two figures are well in line with those independently arrived at from other data.

The *energetics of the sorption process* are in satisfactory agreement with the views developed here. The first amounts of water bound by dry cellulose give rise to a considerable evolution of heat, the differential heat of sorption of the dry material being about 4300 cal/mol H_2O. STAMM and LOUGHBOROUGH [3] have shown the differential heat and the entropy of sorption to be independent upon temperature up to a moisture content of 6% for native cotton fibres, and, moreover, the entropy to be negative.

[1] A. B. D. CASSIE, *Trans. Faraday Soc.*, 41 (1945) 50.

[2] A. J. HAILWOOD and S. HORROBIN, *Trans. Faraday Soc.*, 47B (1946) 84, *Gen. Discussion on Swelling and Shrinking*, held at London, (1946).

[3] A. J. STAMM and W. K. LOUGHBOROUGH, *J. Phys. Chem.*, 39 (1935) 121.

This result affords strong support to the chemical nature of water binding in this region[1] (cf. p. 540). At higher regains, the heat given out becomes dependent on temperature and rapidly decreases according as the moisture content becomes greater, approaching zero at saturation. The observed crossing-over of the sorption isotherms of cotton, measured at different temperatures and at high degrees of saturation, even seems to indicate that the heat effect may become negative near saturation.

STAMM and LOUGHBOROUGH have also pointed out that a striking analogy exists between the thermodynamical behaviour of the cellulose-water system and that of homogeneous binary systems of water with a non-volatile component where exothermic hydrate formation is known to occur, such as the sulphuric acid-water and phosphoric acid-water systems. As early as 1925 KATZ also drew attention to this parallelism[2]. The absorption isotherms of these systems have the same S-shaped form and their thermodynamic characteristics are qualitatively exactly similar. As a matter of fact, the only intrinsic difference between these simpler systems and a macromolecular system like cellulose-water is that the molecules of the initially formed hydrates may freely move in the liquid system and become dissolved in water upon addition of large quantities of this component. In the cellulose system the molecules cannot freely move, but are bound to remain in more or less fixed positions. The swelling towards higher regains may, nevertheless, be considered as a local dissolution of molecular chains, which remain fixed at other places. The dissolved sections will be subjected to BROWNian movements and contribute to an entropy increase.

The analogy is perhaps best shown by plotting the vapour pressure of the two systems cellulose-water and sulphuric-acid water against the mole fraction. (Fig. 35)[3].

Fig. 35. Relative vapour pressure of the systems sulphuric acid-water (□) and cellulose-water (o), plotted against mole fraction.

The system sulphuric acid-water follows the RAOULT-equation at great dilution of the acid, which is demonstrated by the fact that the final slope of the isotherm is the one representing that for ideal solutions (compare p. 514).

In plotting the observations for cellulose the molecular weight of 162 ($C_6H_{10}O_5$) has been employed in the calculations, because the correct value to be substituted is unknown. Neither the value of 162 nor the actual molecular weight of the cellulose would be the correct figure. Some intermediate value, the "effective molecular weight" (with which the partly dissolved chains act) should be taken. Assuming that RAOULT's law is also followed at infinite dilution, its value could, theoretically, be computed by selection of the value of M which would just so modify the curves in Fig. 35 as to exhibit a final slope corresponding to the diagonal of the diagram. The measurements at high relative pressure are, however, not accurate enough for this purpose and, moreover, subject to certain complicating

[1] We may further refer to measurements on the electric conductivity of cellulose fibres as a function of regain by H. MAHLO, *Melliand Textilber.*, 22 (1941) 609. At low moisture contents the resistance is very great, but drops rather abruptly beyond 6—7% regain in cotton and beyond about 10% regain in rayon.

[2] J. R. KATZ, *Ergebn. exakten Naturwiss.*, 3 (1924) 372; 4 (1925) 197.

[3] Cf. J. J. HERMANS, see ref. 2, p. 517.

factors not justifying the application of this procedure. One of these factors is, that infinite dilution is not reached, since we are dealing with a case of limited swelling[1].

The final degree of swelling reached is, here, determined by quite other factors depending on the particular structure of the gel-frame: (The distribution of permanent junction points and the further architecture of the framework cf. p. 572). It is therefore not surprising, that also the "sorption ratio" does not remain a constant in the high vapour pressure range, since the water content of the gel is, in that region, no longer a function of the quantity of amorphous matter, but also of the special structure of the latter.

Summarizing, the sorption of water vapour by cellulose, is, in the low vapour pressure range, a strongly exothermic process in which specific affinities are involved. As the vapour pressure increases, the energy factor gradually becomes less important and the water is bound by less specific and weaker forces. Finally, towards saturation, the entropy factor seems to play the predominant part and the swelling is intrinsically related to a process of dissolution.

The particular characteristics of the water binding at low vapour pressures, gives rise to some very interesting phenomena with regard to the *diffusion of water in cellulose* and to the velocity of conditioning, if cellulose is exposed to atmospheres of different humidity.

Though dry cellulose is to be considered as an extremely hygroscopic substance, the velocity with which equilibrium is attained at low vapour pressures is very small. It is, however, an extremely steep function of the vapour pressure.

The reason of the slow diffusion at low vapour pressures is, that the water molecules in the gel are bound by very strong forces, and can only move from one place to the next after these bonds have been ruptured. At each step, the molecules must overcome a potential barrier of considerable height and this barrier is the higher, the lower the moisture content at which the experiment is carried out.

In the following table some figures are given of the time necessary to reach equilibrium if dry cellulose filaments of about 0.5 mm diameter are exposed to moving air of various moisture content.

TABLE 6

CONDITIONING TIME IN HOURS OF MODEL FILAMENTS IN MOIST AIR

Rel. hum. of air (%)	Time of conditioning (h)
85	5 — 6 h
65	9 — 10 h
35	> 150 h
17.5	> 700 h

For similar reasons it is very difficult to remove the last traces of water from cellulose. The filaments referred to in Table 6 still contained more than 0.5% water after having been kept in vacuo over phosphorus pentoxide at room temperature for 3 months. Without the aid of high temperatures complete drying is practically impossible[2].

The same relationships are responsible for the curious fact that the sorption of water by dry cellulose often takes a heterogeneous course. If dry cellulose fibres

---

[1] In similar systems exhibiting unlimited swelling the analogy is more complete (see the next section).

[2] Cf. the analogous behaviour of gelatine upon drying: S. C. BRADFORD, in J. ALEXANDER'S *Colloid Chemistry I*, New York 1926, p. 764.

are exposed to an atmosphere of, say, 65% rel. humidity and examined under the microscope, a sharp boundary line can be observed to penetrate slowly into the fibre. The same phenomena is observed to occur more rapidly when the fibres are immersed in liquid water. (Fig. 36)[1]. It can be shown that almost immediately after the boundary lines reach the centre of the fibre, moisture equilibrium is attained as well.

The explanation of these phenomena will be clear now: first the outer skin of the object absorbs water. In the layers already moistened, the water can diffuse much more rapidly than in the dry substance. Thus a very steep gradient of moisture content is formed which becomes visible as a sharp boundary line owing to the difference in refractive power of dry and moist cellulose.

Fig. 36. Microphotographs of some successive phases of conditioning upon exposure of dry cellulose filaments to a moist atmosphere.

By measuring the speed of penetration of the boundary line at different temperatures the temperature coefficient of diffusion could be determined. The experiments showed that the velocity of diffusion increased by a factor of 2.2 —2.4 for a temperature interval of 10°. This is a typically. chemical temperature coefficient in conformity with the fact that the first amounts of water are bound by forces of a real chemical nature[2].

Doubtless many other macromolecular systems behave in a similar manner[3]. This also explains why the last few percent of solvents can often only be removed from such systems with extreme difficulty.

Besides the energetic point of view, steric factors will certainly also play their part in these phenomena. At low solvent content, the chains are very densely packed. Solvent molecules, which occur between them, may remain enclosed as in a cage with a grating too narrow to let them pass. Only such molecules will be able to pass from cage to cage whose chemical affinity to the material of the molecular grating enables them occasionally to push their way through by force of chemical interaction. As a matter of fact, molecules, having no chemical affinity for cellulose, as e.g., benzene and ether, when artificially introduced into a swollen cellulose gel, are unable to escape from it on drying, even after prolonged heating at high temperatures. The writer has reported a case of a cellulose filament still containing 64% by weight of ethyl ether after heating it for three hours at 105°[4]. After swelling in water, however, the ether soon comes out as the molecular grating is now widened by the swelling in water.

A few words may be spent here on the systems gelatin water and starch-water whose isotherms have a similar appearance and can be explained in much the same way. The existence of a crystalline hydrate in the system gelatin-water was clearly established by X-ray examination[5]. In the system

---

[1] A more detailed description of these phenomena was given by the author elsewhere (see reference in footnote 3 on page 541).

[2] From this temperature coefficient an activation energy of 13—15.000 cal/mole can be calculated.

[3] Phenomena very much resembling those described here were observed by A. TISELIUS, Z. physik. Chem., A 169 (1934) 425 on the penetration of water into certain minerals belonging to the class of zeolites.

[4] P. H. HERMANS and A. J. DE LEEUW, Kolloid Z., 82 (1938) 63; cf. also R. T. MEASE, Ind. Eng. Chem., 5 (1933) 317 and H. STAUDINGER and W. DÖHLE, J. prakt. Chem., 161 (1943) 219.

[5] K. HERRMANN, O. GERNGROSS and W. ABITZ, Z. physik. Chem., B 10 (1910) 371; J. R. KATZ and J. C. DERKSEN, Rec. trav. chim., 51 (1932) 513.

starch-water, hydrate formation seems to be an intrinsic condition for the formation of a regular lattice order. The crystalline X-ray interferences disappear if the water content of the system is lowered beyond a certain value, indicating a break down of the 3-dimensional order existing in the hydrated state [1].

Finally, we have still to consider the phenomenon of *hysteresis*. To this end, we shall first discuss some phenomena connected with the shrinkage of primary gels formed from cellulose solutions, a subject also dealt with in the preceding section with reference to silicagel.

Also in the case of cellulose, the primary gels have a high watercontent varying with the conditions of preparation, particularly with the concentration of the original solution [2]. Such gels prepared from viscose have a degree of swelling [3], ranging between 6—15 and over. If water is gradually removed by evaporation and the vapour tension of the gel measured as a function of its water content, isotherms of the type schematically shown at *a* in Fig. 37 are observed.

For sake of comparison, the position of the absorption and the desorption isotherms, obtained upon watering and drying, are also shown in the figure at 1 and 2. The primary desorption curve always lies considerably higher than the secondary one.

Dealing with the silicic acid gels (p. 535) we have already stressed the fact that, in cellulose, the shrinkage of the gel continued until all the water is removed. No point comparable to point O in Fig. 31 is observed. Apparently, the framework of the gels is more flexible in cellulose and can, within only a difference of a few per cent., be compressed almost to the density of crystalline cellulose (cf. p. 571).

Fig. 37. Diagram showing the position of the desorption isotherm (*a*) of the primary gel and the absorption (1) and desorption isotherm (2) of the gel after it has once been dried.

A diagram similar to the one shown in Fig. 37 is obtained if the desorption of unripe cotton hairs freshly taken from the boll is examined and compared with their behaviour after having been dried.

The branch *a* in the diagram of Fig. 37 is irreversible and the curves 1 and 2 are only reversible if the complete cycle between O and saturation is considered. Fig. 38 shows which course the isotherms follow if desorption or absorption is interrupted at arbitrary values of the vapour pressure and followed by a reversal of the

---

[1] J. R. KATZ and J. C. DERKSEN, *Z. physik. Chem.*, A 150 (1930) 100; cf. G. CENTOLA, *Atti Acad. Lincei* (6), 14, (1936) 881.

[2] This subject is treated in more detail in the present writers book: *Physics and Chemistry of Cellulose fibres*, Amsterdam 1946.

[3] Degree of swelling = ratio between the volume of the swollen gel to that of the dried one.

process. Every point within the hysteresis loop of the complete cycles is attainable if a suitable procedure of watering or drying is chosen. It will be further clear from this picture that any absorption or desorption isotherm depends up on the condition of the sample taken at its starting point.

All possible curves of a sample which has once been dried lie, however, within the loop between the curves 1 and 2.

The question now arises as to how the phenomenon of hysteresis, which occurs in almost every system consisting of linear macromolecules with a solvent, can be explained.

At equal water content, the vapour pressure is smallest on the desorption curve; the heat and the free energy of hydration are, accordingly, found to be greater than on the absorption curve. On the other hand, accurate measurements in the system cellulose-water have revealed that both the volume of the gel and its refractive index are equal at equal water content in the two cases[1].

It may be assumed that on the desorption branch, where the framework is more and more compressed and the molecular chains come nearer and nearer to each other, more and more cellulose to cellulose contacts are formed in the amorphous regions of the gel. The maximum number of contacts is attained when the gel becomes completely dry. If now, the process is reversed and water added, some of these contacts remain intact. At a given water content, less water is bound directly on the cellulose chains and more looser bound water is present if the absorption branch is considered. According as the water absorption and the swelling increase, more

Fig. 38. Course of the isotherms upon reversal of the cycle at arbitrary vapour pressures.

and more of the temporary cellulose to cellulose contacts are released again. The work necessary to bring about this "uncoupling" is provided at the cost of the free energy of absorption. If the process is now reversed and water removed from the gel, the vapour pressure remains higher, since fewer junction points are present and, consequently, more "free" chain surface is available. The new cellulose to cellulose contacts again gradually formed as a result of the shrinkage, now contribute to the free energy of the process. In other words we may say that the junction point spectrum of the cellulose on the desorption branch is different from that on the absorption branch.

This interpretation of hysteresis is, though cast into somewhat modernized terms, in principle, the same as that earlier proposed by URQUHART[2] and other authors. As stated before, it is by no means confined to cellulose, but has a bearing on a great many macro-molecular systems. It is even not confined to sorption phenomena but applies also to a great many chemical reactions between gels and solutions, if related with swelling.

[1] P. H. HERMANS, loc. cit. (p. 541).

[2] A. R. URQUHART, J. Textile Inst., 20 (1929), T 125. Cf. also W. B. CAMPBELL, Ind. Eng. Chem., 26 (1934), 218. Cf. also G. H. ARGUE and O. MAASS, Canad. J. Research, 12 (1935), 564.

Another interpretation of hysteresis may be given when accounting for the internal tensions occurring in a gel upon swelling and deswelling, as first suggested by CAMPBELL[1]. As a matter of fact, swelling is connected with a deformation of the framework structure and such deformations are always connected with either energy or entropy factors or with both of them. BARKAS[2] has shown that these internal stresses may be quantitatively accounted for formally by an additional either negative or positive hydrostatic pressure of the water in the gel. On this basis CASSIE[3] has endeavoured to explain the hysteresis in the sorption of water by wool. These ideas might be applied to cellulose and other systems as well. We shall, however, not go into the details of this matter, since the way in which the quantitative evaluation of these stresses has been attempted by CASSIE is, as yet, not very convincing. From a qualitative point of view it would seem, however, that the interpretation of hysteresis in the way indicated by BARKAS, is, thus far, the most advanced one.

Hysteresis in the vapour tension isotherms was found in the systems gelatin-water, starch-water, agar-water and many others. It is to be noted that it has been also observed in the equilibria of swollen graphitic acid with water[4], where the water is known to be absorbed in the crystalline lattice of the substance, thus clearly demonstrating, as KATZ[5] justly remarks, that capillary condensation as an explanation is a failure.

δ) *Nitrocellulose.* In this section the system nitrocellulose-acetone will be treated as being an example of sorption gradually leading to unlimited swelling (dissolution), since nitrocellulose is known to be soluble in acetone in all proportions. Thanks to the splendid work of MATHIEU in France and SCHULZ in Germany[6], supplemented by the calorimetric researches of CALVET[7] in the former country, we have a fairly complete and clear picture of the mechanism involved.

Fig. 39 shows the sorption isotherms of the system acetone trinitrocellulose as determined by CALVET at 16°.3 and by DESMAROUX, MATHIEU and PETITPAS[8] at 40°. In Fig. 39a the molar concentration of acetone $n/(n_o + n)$ in the mixture is plotted against the relative vapour pressure p/p_o, in Fig. 39b the 16°.5 isotherm is shown on another scale (weight of acetone per g of nitrocellulose (a) against p/p_o[9]).

In both figures the molar ratios 1 acetone: $1C_6$ and 6 acetone: 1 C_6 are indicated by arrows. Both curves are concave to the pressure axis in the beginning until a molar ratio of about 1 : 1 is reached; this corresponds to a positive heat of sorption. Later the curves bend upwards. In Fig. 39a it will be seen that the curve shows a rather abrupt change in direction at a molar ratio 6 : 1. (The weight scale used in Fig. 39b is not suitable to reproduce this effect).

[1] W. B. CAMPBELL, *Ind. Eng. Chem.*, 26 (1934), 218; cf. also G. H. ARGUE and O. MAASS, *Canad. J. Res.*, 12 (1935), 564.

[2] W. W. BARKAS, *Trans. Faraday Soc.*, 28 (1943), 205.

[3] A B. D. CASSIE, *Trans. Faraday Soc.*, 41 (1945), 50, cf. also p. 89.

[4] J. R. KATZ and J. C. DERKSEN, *Rec. trav. chim.*, 53 (1934), 652.

[5] J. R. KATZ, *Röntgenspektrographie als Untersuchungsmethode*, Berlin-Wien 1934, p. 232.

[6] M. MATHIEU, *La nitration de la cellulose*, Paris 1936; *La gélatination des nitrocelluloses*, Paris 1936; G. V. SCHULZ, *Z. physik. Chem.*, A 184 (1939), 1. (SCHULZ seemed not to know the work of MATHIEU).

[7] E. CALVET et al, *Compt. rend.*, 212 (1941) 542; 213 (1941), 126; 214 (1942) 716, 767; 215 (1942), 138.

[8] J. DESMAROUX, M. MATHIEU et T. PETITPAS, *Compt. rend.*, 213 (1941) 126.

[9] As molecular weight, that of the C_6 unit was taken.

The initial differential heat of sorption of the system is 67.5 cal/g acetone, a figure which is rather modest as compared to the 240 cal/g water in the system cellulose-water. The concavity of the initial part of the curves towards the pressure axis is, accordingly, smaller.

MATHIEU has carefully studied the changes occurring in the X-diagram of various nitrocelluloses (with a nitrogen content ranging from 10.5 to 14%) and could show that the spacing of the 101 planes (the same plane which shifts upon hydrate formation in cellulose) always increases upon sorption of acetone. This plane is the one running
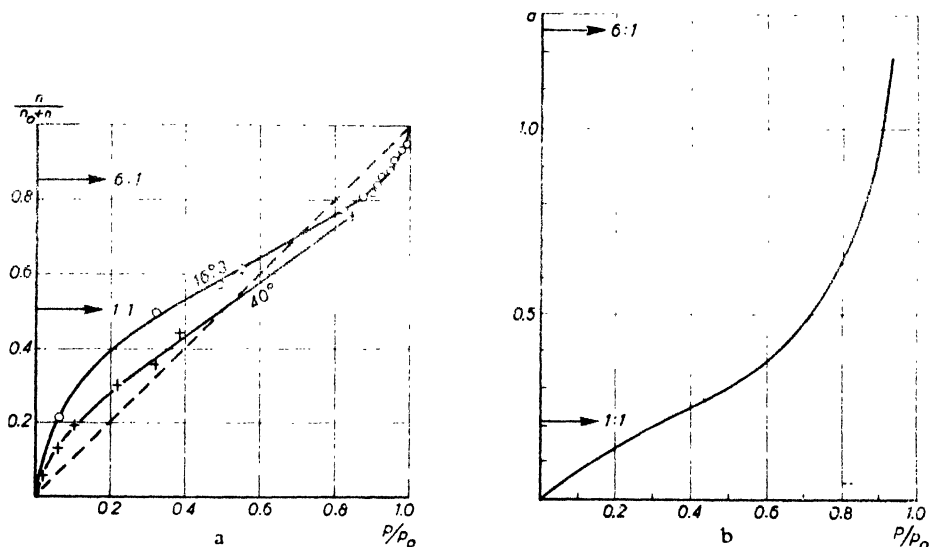
Fig. 39.   Sorption isotherms of the system nitrocellulose-acetone.
(In *a* the molar concentration, in *b* the relative weight of the sorbate is plotted against rel. vapour pressure).

through the hydroxyl groups of the glucose rings in the cellulose lattice and through the nitro groups in that of nitrocellulose. A shift of this plane may be interpreted as due to the insertion of other molecules between the cellulose chains, these other molecules being bound next to the hydroxyls or the nitro groups respectively.

The shift of the 101 interference ceases, when a molecular ratio of 1 acetone molecule per glucose ring is attained. The 002 spacing remains unchanged [1]. So far, the phenomena were identical for all specimen of nitrocellulose studied. If larger quantities of acetone are absorbed, a further shift of the interferences in the X-ray diagram

----

[1] These results are analogous to those obtained by W. TH. ASTBURY in his well known research on keratine. The shifting plane is comparable to ASTBURY's „backbone spacing" (001) which is widened upon reaction with water and strong bases, whereas the "side chain spacing" (002) remains unchanged. These phenomena are specific for many macromolecular chain lattices rather than merely for proteins. It is the same change, too, which is observed upon hydrate formation in gelatin.

is not observed, but they now become gradually blurred, and the crystalline character of the diagram may be said to have disappeared slightly above the molar ratio 2 for dinitrocellulose and at the molar ratio of 2.8 for trinitrocellulose (which contained 2.7 NO_2-groups per C_6). Upon further increasing the acetone concentration, the original character of the diagram rather suddenly disappears. This is accompanied by a change in the appearance of the nitrocellulose itself, which can be observed by the naked eye and was called "gélatination" by MATHIEU[1]. The original fibrous material, first swollen, but still maintaining the fibrous structure and the character of a stiff jelly, now spontaneously flows together and forms a viscous droplet. Whereas the process was reversible until this point (upon lowering the vapour pressure, all previous phases of the X-ray diagram could be reproduced in reverse order), further changes become, from now on, irreversible. The X-ray diagram shows but a broad blurred ring (amorphous halo) approximately corresponding to the former 002 spacing. At further dilution this halo gradually changes into the liquid halo exhibited by acetone.

These observations demonstrate that at the initial phase of sorption aceton molecules penetrate into the lattice and are inserted between the nitro groups. The first molecule being bound, the lattice is widened to a definite degree[2]. Further molecules, up to 1 acetone per nitro group, are then inserted without further change of the lattice dimensions. This stage reached, the lattice structure collapses and regular 3-dimensional order disappears. The system becomes weak and semiliquid, yielding to its surface tension[3]. Upon further addition of acetone "dissolution" begins. As we shall soon explain, more acetone molecules are absorbed with considerable energy and arranged against the chain-molecules, now separating them entirely from each other. If the molecular ration 6 : 1 is reached, another mechanism begins to operate. This occurs beyond a relative vapour pressure of 0.85.

Now, it is of great interest to compare these results of MATHIEU with the calorimetric measurements of CALVET[4]. In Fig. 40 the integral heats of sorption of trinitrocellulose and dinitrocellulose (preparations identical to those used by MATHIEU) are

Fig. 40. Integral heats of sorption in the system nitrocellulose-acetone according to CALVET. Dotted line: trinitrocellulose; full drawn curve: dinitrocellulose. Arrow: heat of dissolution.

---

[1] This is a technical term. It is somewhat confusing that this term is used here for a change from the solid to the liquid state, whereas it is usually connected with the concept of solidification.

[2] That the corresponding interference shifts *gradually* is a phenomenon frequently observed in macromolecular systems. An explanation for this behaviour is given by J. J. HERMANS, *Rec. trav. chim.*, 63 (1944), 211.

[3] A similar change occurs also in the X-ray diagram of the higher paraffins; towards their melting point one of the "side spacings" of the lattice increases more rapidly than the other. The molten paraffin exhibits a broad halo approximately corresponding to the original side spacing.

[4] *Loc. cit.* (p. 548). This work was carried out after the isothermal method, permitting accurate measurements of heat given out over long intervals of time.

plotted against the amount of acetone absorbed expressed in g per g nitrocellulose [1].

For trinitrocellulose the curves could only be followed up to the molar ratio 6 : 1, but for dinitrocellulose (which is more readily soluble) measurements were obtained up to greater dilution. It is seen that a considerable amount of heat is evolved before the molar ratio 6 : 1 is reached. The curves are almost straight ones, indicating that the differential heat of sorption h per mole acetone remains practically constant over this range, every molecule of acetone being bound with equal energy.

At the molar ratio 6 : 1, a total heat of 88.4 cal/g nitrocellulose is evolved. Beyond this ratio no further heat is given out, the integral heat of sorption remaining constant for a while. *It then begins to decrease.* This means that the sorption process after having been *athermal* over a certain range, now becomes *endothermal* upon further dilution of the substance. The free energy of the process remaining positive, this means that the dissolution of the complex: 6 acetone — 1 glucose residue, *now depends entirely on the entropy factor.* In other words, the further dispersion of the chains, their drifting apart from each other, is due to a process of diffusion. The value of the final integral heat of dissolution (measured at infinite dilution) is indicated by a horizontal arrow in Fig. 40. It amounts to only 17.5 cal/g nitrocellulose. Hence 88.4 — 17.5 = 70.9 cal are absorbed in the process of dissolution and are over-compensate dby the entropy gain [2]. (We here further refer to the chapter on Thermodynamics p. 49 and to the chapter on sols without electrolyte character p. 152). We may, hence, consider the complex: 6 acetone — 1 glucose residue, as an addition compound which, upon dissolving in acetone, exhibits a negative heat of dissolution. Investigations of Miss PETITPAS [3] showed that other ketones like cyclopentanone give rise to quite analogous phenomena. Even the lattice widening is identical. She also demonstrated that methyl nitrate, which is not a solvent for nitrocellulose and gives rise to but limited swelling, still shows a similar behaviour.
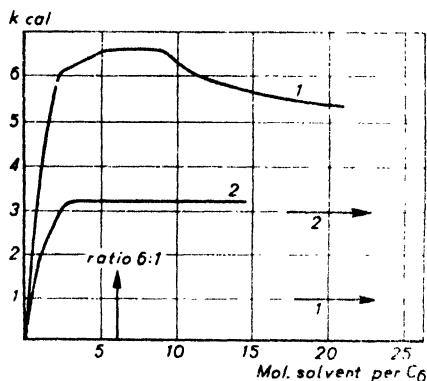
Fig. 41. Integral heats of sorption in the system nitrocellulose-methylnitrate (CALVET). Curve 1 dinitroramie; curve 2 trinitroramie. Arrows: heat evolved upon submersion in the liquid solvent.

[1] The integral heat of sorption H is here defined by $H = \int_{a_s}^{a} h \, da$. At infinite dilution it becomes of course $\int_{0}^{\infty} h \, da$ and in the case of limited swelling $\int_{0}^{a} h \, da$, where a_s is the amount of solvent absorbed at saturation.

[2] This course of the process seems to be rather common in the dissolution of linear polymers, though few examples have been accurately studied. Already as early as 1885 WIEDEMANN and LÜDECKE, *Ann. d. Phys. u. Chemie*, N. F. 25 (1885), 145 reported that the first amounts of water absorbed by gelatin give rise to evolution of heat, but that, upon further dilution, heat is absorbed.

[3] T. PETITPAS, *J. chim. phys.* 37 (1940), 6.

In dinitrocellulose a little more of this solvent was absorbed than in trinitro-cellulose. However, the last phase of the process, that of liquefaction and dissolution, does not occur and the systems maintain their fibrous appearance up to saturation. This may be interpreted thus: the addition complex cellulose nitrate — methyl nitrate is insoluble in methyl nitrate.

The heat evolved upon absorption of 6 moles solvent per glucose residue amounted to 22 cal/g in the case of dinitrocellulose, a figure considerably below that obtained with the ketones. The integral heats of sorption (expressed in mole acetone per mole of the sorbent) measured by CALVET are shown in Fig. 41. The values of the heat evolved upon submersion of the samples in liquid methyl nitrate are indicated with horizontal arrows. (They are comparable, in a sense, to the heat of solution).

The analogy with Fig. 40 is evident and we have here a splendid example of a gradual transition between unlimited and limited swelling, demonstrated with one and the same macromolecular substance.

G. V. SCHULZ[1] has published a thermodynamical study of the system nitro-cellulose-acetone in the range lying between 25% and 99.1% acetone, which appears to be in perfect conformity with the results of the French investigators, and by which our outlook is still further widened. This work will be briefly described here.

The change of the partial free energy $\triangle F_1$ of the solvent, when added to an infinite quantity of a solution, is

$$\triangle F_1 = - \pi v_1 \qquad (12)$$

(see chapter on thermodynamics, p. 84[2]), where v_1 is the partial volume of the solvent in the solution. $\triangle F_1$ can be measured from the osmotic pressure π. In highly concentrated solution this is equivalent to the swelling pressure π_Q:

$$\triangle F_1 = - \pi_Q v_1 \qquad (12a)$$

which, hence, has the same thermodynamical meaning. $\triangle F_1$ can be also determined by measuring the vapour pressure depression of the solvent in the solution:

$$\triangle F_1 = RT \ln \frac{p}{p_o} \qquad (13)$$

In the case of a macromolecular substance, the osmotic measurements can be used in very dilute solutions (compare the Chapter on Molecular Weight Determination, p. 130), whereas the vapour pressure depression equation can be applied at higher concentrations. This was SCHULZ's procedure and, in determining also the change of $\triangle F_1$ with temperature, he could elucidate the thermodynamical characteristics of the system. The change of the internal energy and the entropy of the solvent followed from the equations:

$$\triangle h_1 = \frac{d \left(\triangle F_1 / T \right)}{d(1/T)} \qquad (14)$$

$$\triangle s_1 = - \frac{d \left(\triangle F_1 \right)}{dT} \qquad (15)$$

$$\triangle F_1 = \triangle H_1 - T \triangle S_1 \qquad (16)$$

[1] loc. cit. (p. 548, ref. 6).
[2] The index 1 denotes the solvent component.

Schulz made his investigations with nitrocellulose containing 12.5 — 12.7% nitrogen. (About 2.5 NO_2 per C_6) and with various preparations of molecular weight 50.000 — 443.000, determined with the osmotic method by extrapolation to infinite dilution.

Macromolecular substances generally do not follow Van 't Hoff's law, which
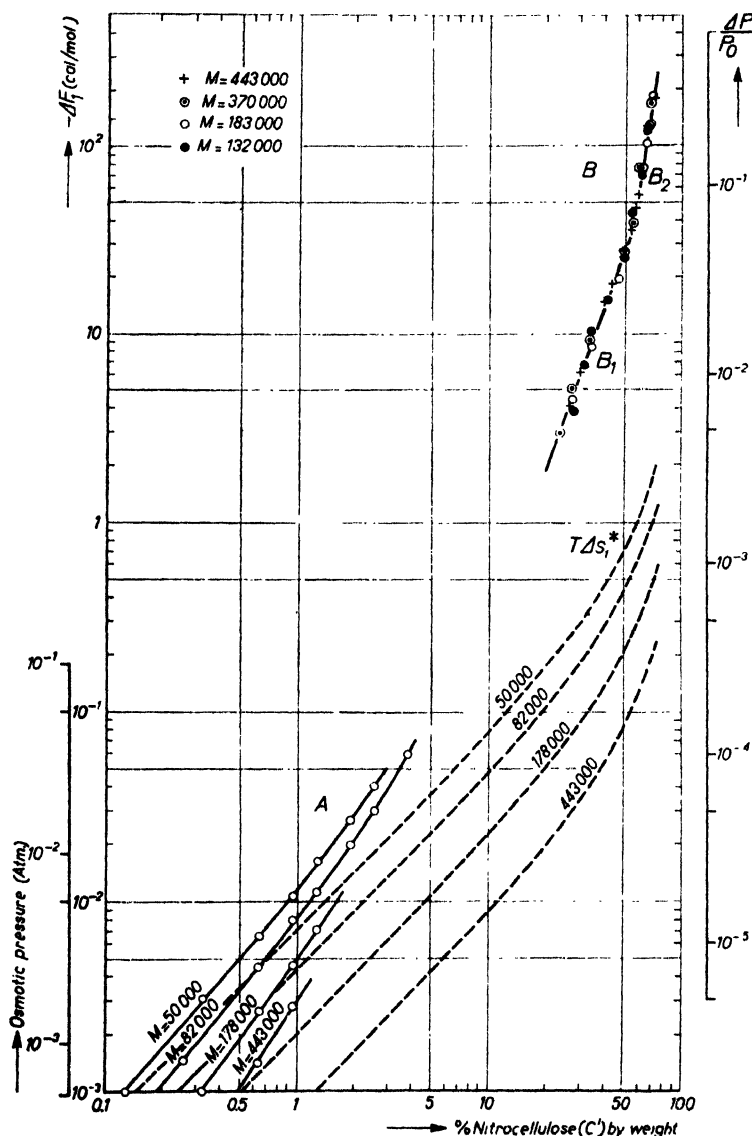
Fig. 42. Schulz's measurements of $\triangle F_1$ (change of free energy of acetone) in the system nitrocellulose-acetone. A. From osmotic pressure, B. from vapour tension depression. The dotted curves correspond to the calculated ideal value of $T\triangle S_1^*$.

only applies if $\triangle H_1 = 0$ and if $\triangle S_1$ corresponds to the ideal value of the mixing entropy $\triangle S_1^*$. At measurable, though still very low concentrations, the osmotic pressure of the samples was considerably greater than the ideal value. This means that either $\triangle H_1 \neq 0$ *or* $\triangle S_1 \neq \triangle S_1^*$ or both.

Now, if $\triangle H_1 < 0$, corresponding to a positive heat effect upon dilution, the osmotic pressure will be higher than the ideal one. We know however from CALVET's work that the heat effect is negative at great dilution and, hence, we may conclude that $\triangle S_1 \gg \triangle S_1^*$, a result which is of great theoretical importance (see chapter on Thermodynamics p. 69). The same was also found in rubber solutions (cf. next section) and has been theoretically elucidated by HUGGINS[1] and FLORY[2]. A high entropy of mixing is inherent to long chain molecules.

Since both $\triangle S_1^*$ and $\triangle S_1$ are functions of the molar fraction[3], determinations of the free energy change $\triangle F_1$ at equal concentration will vary sensibly with molecular weight as long as the entropy term is not negligible compared with the energy term. The latter is, in macromolecular substances, a function of the weight fractions rather than of the molar fractions of solvent and solute, since it depends on the interaction of the total of the monomeric residues with the solvent. If $T \triangle S_1$ becomes small in comparison to $\triangle H_1$, the quantity $\triangle F_1$ and, hence, the osmotic pressure and the swelling pressure become independent of the molecular weight

The outcome of SCHULZ's determinations of $\triangle F_1$ is reproduced by the diagram in Fig. 42. In the region A osmotic pressure measurements were used; p varied from 10^{-3} to 10^{-1} atm. In the region B vapour tension measurements were used down to $p/p_\circ = 0.5 \cdot 10^{-2}$.

In the A region (from $0 - 5\%$ nitrocellulose) $\triangle F_1$ is sensibly dependent of the molecular weight M; in the B region (from $20 - 75\%$) the influence of M falls within the experimental error. From the dotted curve, corresponding to the calculated values of the ideal entropy $T \triangle S_1^*$, it will be easily seen that the contribution of the latter to $\triangle F_1$ is relatively larger in the A region and very small in the B region (less than 1% of the value of $\triangle F_1$).

In the middle of the B-region SCHULZ established a sudden change in the direction of the curve lying at $50-55\%$ nitrocellulose in the mixture. This is obviously the singular composition also stressed by the French investigators and which corresponds to 6 mole acetone per monomeric residue of the cellulosenitrate (cf. Fig. 39a, 40 and 41). SCHULZ concluded that the binding of the solvent below and beyond this point should be of a different kind. We know now from the work of CALVET that at this point the differential heat of sorption changes from a positive into a negative value and from the X-ray work of MATTHIEU, that disintegration of the crystalline regions begins here. This might, with some right, be considered as the point where solution proper begins.

It would seem probable that at higher concentrations than 55% also the differential entropy of the solvent changes sign and becomes negative, since the solvent is now bound by specific forces and the molecules are most probably orientated against the cellulose chains, devoid of their original mobility.

[1] J. M. HUGGINS, *J. Chem. Phys.*, 9 (1941), 440; *J. Phys. Chem.*, 46 (1942), 151; *Ind. Eng. Chem.*, 35 (1943), 216.
[2] P. J. FLORY, *J. Chem. Phys.*, 9 (1941), 660; 10 (1942), 51.
[3] $\triangle S_1^* = RN_2$, where N_2 is the molar fraction of the solute in the solution.

In the majority of the macromolecular swelling systems, (if not in all of them) the rule applies that, in the high concentration range, the heat effect is positive and the entropy effect negative, the free energy, hence, being smaller than the heat effect[1].

By a reasoning not to be referred to here, SCHULZ deduced from his results that the range of the forces between the solvent and the solute decreases less rapidly with the distance than VAN DER WAALS-LONDON attractions would, though the order of magnitude of the forces is equal. He assumed a dipole potential and a polarisation of the molecules of the solvent[2]. The potential of the forces would then be inversely proportional to the 3rd or 4th power of the distance rather than to the 6th power, as for VAN DER WAALS forces. This means that in the system nitrocellulose-acetone, acetone molecules at a distance of 30 Å from the axis of the chains would still be attracted with 0.1 of the force which they undergo when in contact with the chain, whereas e.g. CO_2 molecules in this distance would only exert 0.0005 of the mutual attraction at direct contact. This would mean that even in 5 — 10% solutions nearly all the molecules of the solvent undergo attraction by the chain molecules (Solvation).

A further important conclusion to which SCHULZ directed attention is the following: The consistency of gels and their so-called mechanical properties largely depend on the molecular weight of the macromolecular component. Since the energetic interaction was proved to be independent of the molecular weight in concentrated gels, the molecular attractions can not be considered as being responsible for these effects. They must depend on quite other factors, such as the shape of the molecules or the particular geometric structure of the framework in the gel.

ſ *Some further remarks on swelling systems in general.* The occurrence of an entropy of mixing considerably greater than that calculated for ideal solutions seems to be a general feature of binary systems in which one of the components is a chain molecule. This was shown for esters of fatty acids and long chain alcohols dissolving in solvents like benzene and cyclohexane[3], and seems to apply also for high polymeric substances.

Recently, a very interesting paper on the thermodynamical behaviour of the system rubber-benzene has been published by GEE and TRELOAR[4]. The experiments covered almost the entire range of compositions. Rubber being a polymeric hydrocarbon, it should be expected that no appreciable positive heat effect occurs upon sorption and swelling, since there is no specific affinity between two hydrocarbons. The experiments showed that the process is even an endothermic one from the beginning to the end, a small amount of heat being *absorbed*. The entropy change, however, is considerable and greater than the ideal entropy of mixing.

Qualitatively, the system rubber-benzene behaves exactly like the system consisting of the addition compound of nitro-cellulose with 6 molecules acetone and acetone (see preceeding section).

---

[1] Compare also R. FRICKE and J. LÜCKE, *Z. Elektrochem*, 36, (1930), 309, systems casein-water, agar-water, womans hair-water.

[2] Cf. G. BRIEGLEB, *Zwischenmolekulare Kräfte und Molekülstruktur*, Stuttgart 1937.

[3] J. N. BRÖNSTED and P. COLMANT, *Z. physik. Chem.* 168 (1934), 381; K. H. MEYER and R. LÜHDERMANN, *Helv. chim. acta*, 18 (1935), 307; CH. G. BOISSONNAS, *Helv. chim. acta*, 20 (1937), 768.

[4] G. GEE and L. R. G. TRELOAR, *Trans. Faraday Soc.*, 38 (1942), 147; G. GEE, *ibid.* 38 (1942). 276, 418. (Cf. the chapter on Sols without electrolyte character p. 158).

It will be clear that, when dealing with sorption and swelling in linear polymers, we may in general distinguish two phases. Firstly a chemical interaction with the solvent may occur. This depends up on the chemical nature of the polymer and that of the solvent and their mutual specific affinities. Such a chemical interaction involves a positive heat effect, and one or more addition compounds may be formed. Often, the formation of such compounds can be detected by characteristic changes in the X-ray diagram of the original substance; a new more or less well defined diagram, that of the compound, appears. This is the case if the compound formation extends itself to the crystalline portion of the high polymer. Absence of a new diagram does, however, not imply that no formation of addition compounds takes place, since it may occur that compound formation is confined to the amorphous portion exclusively.

Secondly we have the process of dissolution proper. It is due to the entropy gain upon mixing. In the few cases investigated so far, it is accompanied by an absorption of heat.

b. Swelling

Typical swelling is an exclusive property of macromolecular systems, though not every macromolecular system exhibits the phenomenon of swelling. The conditions at issue have been treated in Chapter II and Chapter VI (HOUWINK) and have, moreover, been recapitulated earlier in this Chapter on p. 512. The mechanism of swelling has also been dealt with in Chapter III and was referred to in the section on sorption. In a sense, we may say that swelling is a visual demonstration of the sluggish mobility (low diffusion velocity) of large molecules. Small molecules more easily follow their diffusion tendency and, provided energetic conditions do not oppose, they easily and rapidly diffuse into and disperse themselves throughout the solvent. Though the *final* entropy gain upon dissolution is, as we have seen, greater in large chainlike molecules than in small ones, the *velocity* of the dissolution process is smaller and its final state is more slowly reached. Swelling may be considered as an initial transitionary phase of dissolution. That it may stop before complete solution is reached (limited swelling), is a more or less incidental circumstance connected with the particular structure of the solute as a system with a predetermined permanent coherence (p. 165) or with particular phase equilibrium conditions in the binary system concerned (cf. p. 87).

THOMAS GRAHAM defined colloids as substances with small diffusing power. Swelling, which is considered as being a typically „colloidal phenomenon" seems to be actually due to this very property. It is interesting, that macromolecular substances with spherical molecules do not exhibit swelling. A typical example is glycogen, which according to H. STAUDINGER consists of globular molecules. Even samples with a molecular weight of 800.000 do not swell upon dissolution. This demonstrates that slow diffusion cannot be considered as the sole factor controlling swelling.

It would seem that an explanation is possible along the following lines: The cohesive forces per unit weight of substance are of the same order of magnitude in chain molecules and their monomers. Per molecule, however, the cohesion is, of course, far larger for the macromolecules, a fact which explains many characteristic macromolecular properties. However, considering spherical macromolecules having a much smaller relative surface per molecule, both the cohesion per unit weight and per molecule will be smaller than in linear molecules. This energetic factor will facilitate their dissolution. Moreover, the moment a spherical molecule is freed from

the solid, it will at once follow the resulting velocity imposed upon it by the collisions with the solvent molecules and move away as a whole. In linear molecules, the impulses of the collisions have another effect in that the resultants of thermal motion may have different directions at various parts of the chain. Moreover, the linear molecule may be more tightly bound at one spot than at another one; the spherical molecule will be less liable to these factors. In the following sections we shall discuss some further aspects of swelling.

b. 1. Swelling pressure

If a swelling substance is tightly enclosed in a vessel with a wall permeable to a swelling solvent, and the latter is allowed to diffuse into the vessel, the dilation tendency of the gel may give rise to a pressure.

In the diagrammatic figure 43, A represents the swelling substance and B a piston permeable to the solvent S. To prevent the piston from rising, a pressure P per unit surface must be excerted on it. This pressure, just counterbalancing swelling, depends upon the initial degree of swelling of the gel[1]. If this pressure is raised beyond this value, solvent can be pressed out of the gel.

Physically, the swelling pressure of a gel may be considered from exactly the same point of view as the osmotic pressure of a solution (see p. 83) and may be expressed in terms of the same thermodynamical quantities. The work done by the solution or by the gel against this pressure is equal to the free energy decrease of the system. In both cases the pressure may be interpreted as the hydrostatic pressure which has to be applied to the solution or to the gel in order to increase

Fig. 43.
Swelling pressure.

the vapour pressure of the solvent in the solution until it becomes equal to that of the pure solvent. Kinetically the pressures may be regarded as due to the tendency of solvent molecules to flow from the solvent to the solution as a result of the free energy decrease involved.

According to experiments of POSNJAK[2] with rubber gels, the swelling pressure π may be represented by an equation of the form

$$\pi = kc^n$$

where c is the concentration of the rubber in the gel and k and n are characteristic constants lying usually between 2 and 3. This must be regarded as an empirical equation having no theoretical background whatever[3].

b. 2. The degree of swelling

α. *Systems without electrolyte character.* A macromolecular substance, when immersed in a non dissolving but swelling liquid, shows the phenomenon of limited swelling. A more or less well defined equilibrium size is reached. For a given gel, the final degree of swelling depends on the nature and the composition of the solvent.

---
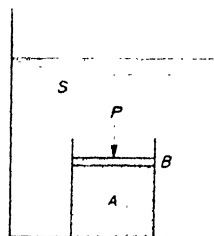
[1] The pressures involved may reach very high values. This pressure has been utilized in practice particularly in ancient times before explosives where available. Rocks were split by hammering dried wood into a hole and subsequently moistening the latter with water.

[2] E. POSNJAK, *Kolloidchem. Beih.*, 3 (1912), 417.

[3] Cf. also footnote 2 on page 524.

Usually, any degree of swelling can be imposed on the gel by selecting the proper solvent or solvent mixture. Gradual transitions exist between limited and unlimited swelling. The exact theoretical treatment of the subject is a difficult and complicated one. Very valuable contributions on a thermodynamical basis, dealing with the swelling of rubber in various liquids, have recently been published by GEE and TRELOAR[1]. The thermodynamic criterion for equilibrium swelling is given by the condition $\triangle \mu_o = \triangle h_o - T \triangle s_o = 0$, where $\triangle \mu_o$ represent the GIBBS free energy of dilution,) $\triangle h_o$ the heat of dilution and $\triangle s_o$ the entropy of dilution. In many cases the heat of dilution can be approximated by VAN LAAR's equation (see p. 64 in the chapter on thermodynamics).

$$\triangle h_o = k_o v_r{}^2 \tag{17}$$

where v_r is the partial volume of the solid in the gel and k_o is a constant. It is obvious that v_r is related to the degree of swelling by the equation[2]

$$q = 1/v_r . \tag{18}$$

From (17) it follows that the equilibrium condition becomes

$$T \triangle s_o = k_o v^2{}_r \text{ or } v_r = \sqrt{T \triangle s_o/k_o} \tag{18}$$

The constant k_o is characteristic for the solvent and can often be represented fairly accurately by an equation borrowed from HILDEBRAND et al[3]

$$k_o = a V_o \left[\sqrt{E_1/V_1} - \sqrt{E_2/V_2}\right]^2 \tag{19}$$

E_1 and E_2 being the molar cohesive energies of the solvent and of rubber respectively and V_1 and V_2 their molar volumes. a is a constant whose order of magnitude is 1. The cohesive energy for rubber was found by GEE to be 66 cal/cm³ and it could be shown that, using the values of k_o given by (19) and GEE's estimate of the entropy of swelling in rubber, results consistent with the swelling data for rubber as collected by WHITBY et al[4] were obtained.

It is evident that swelling will go to its maximum extent if k_o becomes zero, i. e if $E_1/V_1 = E_2/V_2$. Hence, this will occur for rubber in a solvent whose cohesive energy is 66 cal/cm³. Values of E_1/V_1 either higher or lower than 66 cal/cm³ will yield positive values of k_o and therefore lower the degree of swelling q. Hence, if q is plotted as a function of $x = V V_1 \left[\sqrt{E_1/V_1} - \sqrt{66}\right]$ one should obtain a curve approximating in form to a GAUSSIAN error function, with its maximum at $x = 0$. Gee[5] has shown this to be roughly true with certain well-defined exceptions. In a later paper GEE[6] has treated the swelling of rubber in mixtures of two solvents and succeeded in developing a theory reasonably consistent with experimental data. In some pairs of solvents swelling is simply additive, i. e. a linear function of the

[1] G. GEE, *Trans. Faraday Soc.*, 38 (1942), 276, 419; 40 (1944), 468; J. FERRY, G. GEE and L. R. G. TRELOAR, *ibid.*, 41 (1945), 340.
[2] GEE defines the swelling Q as $(1 - v_r)/v_r$. In this volume we shall consider as degree of swelling the ratio of the volume the swollen gel to that of the dry substance, denoted by q.
[3] J. H. HILDEBRAND and G. SCATCHARD, *Chem. Reviews*, 8 (1931), 321; *J. Am. Chem. Soc.*, 56 (1934), 995.
[4] G. S. WHITBY, A. B. A. EVANS and D. S. PASTERNACH, *Trans. Faraday Soc.*, 38 (1942) 269.
[5] G. GEE, *Trans. Faraday Soc.*, 38 (1942), 418.
[6] G. GEE, *Trans. Faraday Soc.*, 40 (1944), 468.

solvent composition. In others, curves either convex or concave to the axis representing the solvent composition are obtained. Also curves with a definite maximum or minimum may occur. Fig. 44 (borrowed from GEE) shows the swelling of Buna S in various mixtures.

Curves 1 and 4 are particularly interesting. Curve 1 shows that the swelling power of n-butyl acetate is slightly reduced by adding a small proportion of the much more powerful swelling agent chloroform. Curve 2 represents an example of a pair of liquids which, in combination, give rise to enhanced swelling. From GEE's theory it could be predicted that a mixture of liquids will give rise to enhanced swelling if the cohesive densities EV of the two liquids lie on opposite sides of that of the polymer (one being greater and the other smaller).

The dissolving power of solvent mixtures follows the same general rules and a mixture of properly selected liquids, which each by themselves merely give rise to limited swelling, may be a solvent. Numerous examples were already well known from practice[1]. In such cases two critical solubility limits will exist.

In a great many cases the limiting condition for the applicability of GEE's theory (no marked specific energetic interactions between the solvent and the solute) is not fulfilled. It would seem, however, that, from a theoretical point of view, such cases may be reduced to the former by considering the exothermic addition compounds formed by

Fig. 44. Swelling of vulcanised Buna S in mixed liquids.
1. Chloroform — n butyl acetate; 2, hexane — ethyl acetate; 3. ethyl ether — n propyl acetate; 4. ethyl methyl ketone — ethyl acetate.

the macromolecular substance as the solute instead of the dry substance (cf. Section 6a. 3δ, p. 548), thus separating the process of swelling in two consecutive phases, a "chemical" and a "physical" one.

The influence of temperature on swelling can be readily understood from the thermal behaviour of the system concerned. If swelling is an exothermic process, the equilibrium degree of swelling will be lowered by a raise in temperature and vice versa. Since intramicellar swelling is always an exothermic process, it has to be expected that the corresponding shift of the lattice spacings which can be read from the X-ray diffraction photographs, decreases with rising temperature. This has been experimentally verified for the intramicellar swelling of graphitic acid and collagen in water bij DERKSEN[2]. The swelling of cellulose in water and sodium hydroxide solutions, which represents an exothermic process in all phases of the swelling, is actually enhanced by lowering the temperature. In the foregoing (Section a. 3 δ and ε) we have seen that the subsequent intermicellar swelling in a great many cases is an endothermic process. In this phase the swelling will, consequently, be enhanced

---

[1] E. g. cellulose nitrate in alcohol + ether (collodion), polystyrene in acetone + n-propyl laurate (BRÖNSTED and VOLQUARZ, *Trans. Faraday Soc.*, 36 (1940), 619). A related phenomenon is the increased ease of solution of rubber in hexane or partially etherified cellulose-derivatives in hydrocarbons on adding alcohols or ketones in a small proportion.

[2] J. R. KATZ and J. C. DERKSEN, *Rec. trav. chim.*, 53 (1934), 652; J. C. DERKSEN, *Thesis*, Amsterdam 1935, p. 19 and 35.

by a temperature rise. The swelling degree of gelatin gels increases with temperature though the swelling of its crystallites is simultaneously reduced. It will be hardly necessary to say, that the *velocity* of swelling (the speed with which the equilibrium condition is reached) always increases with the temperature.

β *Systems with electrolyte character.* The phenomena concerned with the swelling of "hydrophilic" macromolecular substances, such as proteins and cellulose, in aqueous solutions of electrolytes are more complicated and more difficult to understand theoretically, particularly if the macromolecules themselves represent electrolytes too, i. e. if they bear typically ionisable groups as e. g. the proteins. Since a great deal of the subject-matter has been treated elsewhere in this book (see Chapter VII), we shall confine ourselves here to some of its main features only.

Thus far, the theoretical considerations dealing with this kind of systems have been of quite another nature than those tentatively advanced to explain the behaviour of the systems without electrolyte character. Obviously, a certain gap has remained between the two domains of thought. Nevertheless it would seem that in the near future this gap may be closed.

The swelling of macromolecular systems with electrolyte character in pure water may be dealt with in the same way as the cases considered in the preceding section. If electrolytes are added, a new situation emerges. To handle it, the theories of common electrolytes must be brought into the discussion. To a certain extent the phenomena involved may, however, be considered as being analogous to those concerning the swelling of systems without electrolyte character in solvent mixtures of various composition. It should be possible, at least, to apply similar thermodynamical reasonings in both cases. In the latter, however, the thermodynamical quantities depend up on different factors. As far as known to the author, a consistent theoretical treatment along such lines has thus far not been given. Since a more general theoretical base is lacking, we shall content ourselves with a brief survey of some typical phenomena, against the (necessarily somewhat one-sided) theoretical background thus far developed.

We have seen that the tendency of gels to swell may be expressed in terms of osmotic or swelling pressure, these quantities actually representing a measure of the tendency of the solvent to enter into the gel, which here may be compared to a solution (Cf. p. 543). If ionisable systems are concerned, the so-called DONNAN equilibria should be taken into account[1].

The elementary theory of the DONNAN equilibria has been treated in Part I, Chapter X. Considering a gel of a macromolecular electrolyte in equilibrium with a solution, a situation is met with similar to that of the case of an electrolyte with one ion which cannot diffuse through the membrane. The macromolecules are bound to remain in the gel since they form an intrinsic part of the framework. The concentration of the added electrolyte will always be different in and outside the gel.

For simplicity, let us assume that a unit volume of a gel (1) consisting of MCl swollen in water is immersed in an equal volume of KCl solution (2). The initial molar concentrations of MCl and KCl be respectively c_1 and c_2.

---

[1] See for a detailed survey T. R. BOLAM, *The Donnan equilibria*, London 1932.

gel		solution	
c_1	M^+	c_2	K^+
c_1	Cl^-	c_2	Cl^-

gel		solution	
c_1	M^+	c_2-x	K^+
c_1+x	Cl^-	c_2-x	Cl^-
x	K^+		

       initial condition   equilibrium condition

Let equilibrium be attained when x mole KCl have diffused into the gel. Hence,

$$(c_1 + x)x = (c_2 - x)^2 \qquad (20)$$

$$\text{or} \qquad x = \frac{c_2^2}{c_1 + 2c_2} \qquad (21)$$

If c_1 is small as compared to the KCl concentration c_2, approximately half of the KCl will diffuse into the gel ($x = \frac{1}{2}$) and its concentration is almost equal in and outside the gel. If, on the contrary, c_1 is very much greater than c_2, practically all the K-ions will remain outside the gel. If $c_1 = c_2$, one third of the KCl goes into the gel and two thirds remain outside.

In an analogous way the equilibrium condition can be found when the gel and the solution have no common ion, or when one of the ions has a valency differing from unity. Further, the theory of strong electrolytes may be introduced, substituting ionic activities for ionic concentrations[1].

As to swelling, the following reasoning was put forward. The osmotic pressure of the liquid outside and inside the gel will be determined by the total concentration of the mobile ions. In our example, these concentrations in equilibrium condition are $c_1 + 2x$ inside and $2(c_2-x)$ outside the gel. The difference between these two values $c_1 + 4x - c_2$ would determine the degree of swelling and also represent the swelling pressure.

In order to form ourselves a clear picture of the swelling process, the idea that it is due to some kind of "internal pressure" seems not to be the most adequate conception[2]. We should rather say that the free energies of the solvent inside and outside the gel are different and that, hence, the solvent has the tendency to pass from the phase where it has the highest to the phase where it has the lowest free energy.

Upon verification under such conditions, the DONNAN theory, as applied to swelling, held satisfactorily in various instances. It will be clear that this simple theory cannot bear a general character and that its applicability will be confined to those cases where its premises may be expected to be consistent i. e. in the swelling and contraction of relatively diluted gels upon changes of the electrolyte concentration in dilute equilibrium solutions[3].

---

[1] J. LOEB, *J. Gen. Physiol.*, 3 (1921), 85, 547, 667, 691; *J. Am. Chem. Soc.*, 44 (1922), 1930.

[2] A block of metal, when heated, expands. If the dilatation is counteracted by enclosing the metal in a vessel with a lower coefficient of dilatation, a large pressure will be exerted on the vessel. In the swelling process there is no more reason to speak of an internal pressure causing swelling, as there is to speak of an internal pressure in the metal block, as a result of heating. The expansive tendency can in both cases be *measured* in terms of a pressure, but it would be too simple a picture to say that expansion is caused by an internal pressure.

[3] The swelling of dry gelatin e. g. will be largely governed by the factors treated of in section a.

Under such conditions the theory held satisfactorily in various instances. LOEB, e. g., has shown that the osmotic pressure of a dilute protein solution in dependence on the hydrogen ion concentration, which shows a maximum value at a definite pH, satisfactorily conforms with the theory. If gelatin is allowed to swell in water and then submerged in dilute hydrochloric acid, the degree of swelling reaches a maximum at about the same pH.

PROCTER[1] was the first to show that these phenomena could be explained if due allowance was made for the DONNAN equilibrium between the gel and the liquid, and if the formation of ionisable salts between gelatin, acting as a weak base, and hydrochloric acid was assumed. The equilibrium is represented by the following diagram

z	GH^+			
z	H^+		x	H^+
$y + z$	Cl^-		x	Cl^-
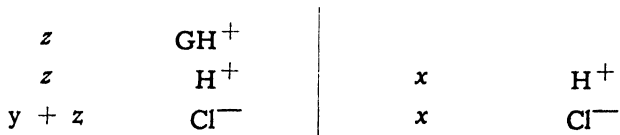
The ions GH^+ are the non diffusible ones. (Their nature is comparable with that of the ammonium ion). According to equation (20) the equilibrium condition will be

$$x^2 = y(y + z)$$

Determining the H-ion concentration inside and outside the gel at various concentrations of acid in the liquid, it could be shown that the theory was in conformity with the facts. PROCTER[2] held the view that the swelling of the gel is due to the difference in osmotic pressure of the solution in- and outside the gel, resulting from the different concentration of mobile ions. The distending force, hence, would be an osmotic one. The opposing force, making equilibrium with the former, would be the cohesion of the gel framework.

From electric potential measurements LOEB[3] could derive the pH of the liquid. The results were in conformity with the DONNAN equilibrium. From the experimental data the osmotic pressure difference $\triangle \pi$ could be calculated. Plotted against the pH of the outer liquid, the maximum value of $\triangle \pi$ coincided with the pH corresponding to maximum swelling, as shown by the following selected figures.

TABLE 7

OSMOTIC PRESSURE DIFFERENCE $\triangle \pi$ IN DEPENDENCE ON THE PH OF THE EQUILIBRIUM SOLUTION OUTSIDE THE GELATIN GEL.

PH	3.89	3.04	2.65	2.44	2.16	1.95	1.82	1.49
$\triangle \pi$	0.4	3.1	5.0	4.9	2.4	1.5	1.6	1.7

The essential point is that starting from the iso-electric point (which lies at about 4.5 in gelatin) and lowering the pH, the charge of the gelatin increases but also

[1] H. R. PROCTER, *J. Chem. Soc. London*, 105 (1914), 313.
[2] H. R. PROCTER and J. A. WILSON, *J. Chem. Soc. London*, 109 (1916), 307.
[3] R. F. LOEB, *J. Gen. Physiol.* 4 (1921—1922), 351.

the total concentration of the ions in the gel. According to DONNAN the osmotic pressure will first increase owing to the higher charge and then diminish owing to the increased concentration of the ions.

Similar results were reported by various other authors[1]. Values of pH $\leq$ 2.5 in the outer liquid and $p_H \leq$ 3 in the gel were invariably obtained.

PROCTER and BURTON[2] showed that the volume increase of a given quantity of gelatin, when swelling in hydrochloric acid of various concentrations, is proportional to the osmotic pressure difference $\triangle \pi$ [3]. By an extension of the theory WILSON and WILSON calculated the degree of swelling in dependence of the H-ion concentration and found figures in conformity with experiments.

LOEB and KUNITZ[4] showed that, at a given pH, dibasic acids gave rise to a lower degree of swelling than monobasic ones, just as required by theory. Some data are reproduced in Fig. 45. It is seen that the swelling is independent of the nature of the acid and is merely determined by the pH and the valency of the acid anion. The small departure shown by acetic acid is due to the high concentration of this acid necessary to attain pH = 3. The medium water is thereby modified to such an extent that solvent effects of a different nature may be expected. The influence of additions of neutral salts could be also satisfactorily accounted for on the basis of the theory [5].

Fig. 45. Swelling of gelatin in dependence of hydrogen ion concentration (according to LOEB and KUNITZ) 1. Acetid acid; 2. HCl, HBr, HI, HNO₃, lactic acid; 3. H₂SO₄, sulpho salicylic acid.

It would seem that the osmotic nature of the swelling of gelatin, as affected by acids, is rather well established [6]. This is not surprising, since the swelling of gels without electrolyte character may be also expressed in terms of osmotic pressures (see Section a), though this terminology is not a very attractive one. The effects of ionic concentration can be expressed in terms of the proper thermodynamic quantities as well, and this more adequate procedure will lead to the same results.

Obviously, the intrinsic nature of the swelling phenomena considered here is the diffusion tendency, or, in other words, the tendency of the system to reach a state of higher dilution, which is aided by the presence of mobile ions in the cases just considered.

The point of view put forward by PROCTER as early as 1914 that, in as far as limited swelling is concerned, this tendency of the gel to expand is counterbalanced by the "cohesion forces of the framework" seems to be a particularly important one.

[1] e. g. D. JORDAN LLOYD, *Biochem. J.*, 14 (1920), 147, 584; 24 (1930), 1460, NORTHROP and KUNITZ, *J. Gen. Physiol.*, 12 (1928—1929), 537.

[2] H. R. PROCTER and BURTON, *J. Soc. Chem. Ind.*, 35 (1916), 404.

[3] Cf. J. A. and W. H. WILSON, *J. Am. Chem. Soc.*, 40 (1918), 886.

[4] R. F. LOEB and M. KUNITZ, *J. Gen. Physiol.*, 5 (1922—1923), 665.

[5] R. F. LOEB; R. F. LOEB and M. KUNITZ, *loc. cit.*

[6] Modifications of the theory as proposed by D. JORDAN LLOYD (*loc. cit.*), do not conflict with the basic concepts. Other theories assuming an influence of hydration (PAULI) or mutual repulsion of electric charges on the framework of the gel (TOLMAN) have been proved untenable by experiments of S. GOSH, *J. chem. Soc.*, London 1928, p. 711.

From recent work the idea has emerged that swelling and contraction of macromolecular gels may be connected with a stretching and folding of the molecular chains between the junction points and that similar changes of the form of the molecular chains occur in deformation of gels (cf. Section 6 b. 3, p. 567). The cohesive tension then would be intimately connected with the resistance of the chains against changes of their form. This resistance may involve either an entropy or an energy effect or possibly both (cf. Section b 3).

Full credit should be given to PROCTER for the vision, quite correctly deduced from his observations on gelatin gels, that these objects should be considered as exhibiting the properties of true solutions and as containing a framework of molecular dimensions. At a time when scores of intrinsically different theories on the colloidal structure of such objects were launched, often full of undigested and unclear concepts, or hypotheses lacking a sound basis, which for a great many years to come haunted the minds of colloid chemists, this was doubtless a great achievement.

NEALE [1] endeavoured to explain the swelling of cellulose in sodium hydroxide solutions (which shows a marked maximum at a concentration of about 2 mole/liter) on the basis of the DONNAN theory, thereby considering cellulose as a weak acid capable of splitting off hydrogen ions. Though the results were consistent with the experimental data, the behaviour of this system seems to require another explanation, since HESS and coworkers [2] have shown that the results of experiments on the distribution of added neutral salts lead to the conclusion that the theory does not hold. Moreover, the assumption that cellulose acts as a weak acid seems not to be sufficiently justified. It is true that carboxylic endgroups frequently occur in cellulose, but these cannot be held responsible for the acid character postulated by NEALE, since they have a much larger dissociation constant than the one required by NEALE's theory.

It would be of interest to try a quantitative theory, based on similar ideas as those which have been developed by GEE for the swelling of rubber in mixtures of two solvents (cf. Section 6, b. 2a). The solutions of sodium hydroxide of various concentrations should thereby be regarded as solvent mixtures of different composition and their thermodynamical data studied.

The influence of neutral salts in moderate concentrations represents quite another chapter. It has been recognized long since that the action of salts with the same kation but different anion, as well as salts with the same anion and different kation, may be arranged in a certain order termed HOFMEISTER series or *"lyotropic order"*.

$$SO_4 - acetate - Cl - NO_3 - CNS - I$$
$$Mg - NH_4 - K - Na - Li.$$

Swelling is enhanced from left to right in this series. Thus, the salts $MgSO_4$ and LiI represent those with a maximum difference in their effect.

The same or almost the same succession was found to hold in various instances where the influence of salts on "lyophilic" macromolecular systems was studied, as e. g. the viscosity of starch sols, the flocculation of agar agar and gelatin solutions, the melting point and the elasticity of gelatin gels. Also in other cases where water is an active component of a system similar rules were found (influence on the viscosity of sugar solutions, on the mutarotation of glucose, on the hydrolysis of esters by bases, and so on). The vapour pressure and freezing point depressions of the salts follow the same order.

Qualitatively speaking, it can be said that the "activity" of the water, its properties

---

[1] S. M. NEALE, *Shirley Inst. Mem.*, 8 (1929), 87; 9 (1930), 21; *J. Textile Inst.*, 20 (1929), T 373; 21 (1930), T 225.

[2] K. HESS, C. TROGUS and O. SCHWARZKOPFF, *Z. physik. Chem.*, A 162 (1932), 187.

as a solvent or as a reactant are affected always in much the same way by the addition of a given neutral salt. This leads to the conclusion that it is the interaction of the salt with the water which is the predominating factor responsible for these phenomen. Though an entirely satisfactory quantitative theory does not exist, it would seem that, broadly speaking, a general insight in the subject is at present within our reach.

In the dissolution of a salt in water, the very considerable lattice energy of the salt must be overcome. The work necessary to separate the ions is, however, compensated by the hydration energy of the ions, which is, as a rule, of the same order of magnitude. So it is conceivable that the heat of dissolution of salts like $NaCl$ is only slightly negative. In a number of cases it is even positive. The energy of hydration then surpasses the work necessary to overcome lattice cohesion. In insoluble salts like CaF_2 and $BaSO_4$ the lattice energy is so large as compared with that of ionic hydration that the entropy gain of dissolution is no more sufficient to bring about dissolution. The positive energy of ionic hydration is demonstrated by the existence of many crystalline hydrates. Water can penetrate into the lattice of the anhydrous salt with evolution of heat. This hydrate formation may be considered as a preliminary phase of dissolution. According to the nature of the ions and the number of water molecules previously bound as water of crystallisation, the heat of dissolution of the anhydrous or the hydrated salt will be positive or negative. Roughly speaking, salts not capable of forming a crystalline hydrate will dissolve with absorption of heat and vice versa. Most crystalline hydrates containing the maximum amount of hydrate water dissolve with a negative heat effect.

The heat absorbed upon dissolution of salts with univalent ions, not capable of hydrate formation, will increase according as the ionic radii of the ions become greater, since the energy of hydration decreases in the same sense[1]. Upon dissolution of KI and NaCl more heat per gram equivalent is absorbed in the former case. Sometimes the differential heat of dissolution of the anhydrous salt is positive in concentrated solutions and becomes negative upon further dilution. This will occur if the salt forms a crystalline hydrate, which then dissolves with absorption of heat[2].

All these phenomena can be qualitatively understood, when allowing for the elementary thermodynamical principles and they are quite analogous, to those elsewhere discussed in this volume, dealing with the interaction of macromolecular substances with a solvent.

BUCHNER and coworkers[3] established certain simple quantitative relationships in the action of salts on colloidal systems according to the HOFMEISTER rule. To a series of univalent anions they could even ascribe a numerical figure, termed "lyotropic value", standing for the quantitative location of the ion concerned in the HOFMEISTER series. These figures were found to yield a linear relationship with the reciprocals of the ionic radii[4].

Roughly speaking, we may also state that the differential heat of dilution of a salt for the concentrations at which its "lyotropic" effect is considered, becomes

---

[1] In the case of multivalent ions, the charge of the ion interferes as an other factor.
[2] This is e. g. found for sodium hydroxide.
[3] E. H. BUCHNER and D. KLEYN, *Proc. Acad. Sci.* Amsterdam, 30 (1927), 740; E. H. BUCHNER and G. POSTMA, *ibid.*, 34 (1931), 699; E. M. BRUYNS, *ibid.*, 35 (1932), 107; cf. also J. H. C. MERCKEL, *Thesis*, Amsterdam 1934; *Kolloid-Z.* 78 (1937), 41, 339; J. H. C. MERCKEL and E. H. WIEBENGA, *Kolloid-Z.*, 80 (1937), 315; J. L. OUWELTJES, *Thesis*, Amsterdam 1942.
[4] E. H. BUCHNER, A. VOET and E. M. BRUINS, *Proc. Acad. Sci. Amsterdam*, 35 (1932), 763.

smaller, or shifts towards higher negative values respectively, according as its place in the HOFMEISTER series shifts to the right. This means that the internal energy of the solution is enhanced upon dilution. The fact that, nevertheless, the free energy of dilution decreases upon dilution, is due to the still greater entropy gain of the process.

The "lyotropic action", which we shall define as increasing from left to right in the HOFMEISTER sequence as reproduced above, may be broadly embodied in the statement that it enhances the solubility of the colloid. (Solutions are less readily flocculated, the "melting point" of gelatine gel is lowered, their extensibility for a given load enhanced, the gelatination temperature of the solution depressed, the degree of swelling increased a.s.o.).

We may now venture to picture lyotropic action as a matter of competition between the dissolved salt and the macromolecular component towards the water. Let us assume that the entropy of dilution of a series of simple salts is practically equal. If a macromolecular system interacts with the salt solution — let us e. g. consider a junction point between two adjacent chains as occurring in a gelatin gel (see diagram in Fig. 46) the question as to whether the junction point will be loosened or tightened by the action of the solution depends again on the free energy change involved. Let the partial thermodynamic quantities referring to the junction points and to the salt solution be denoted by the subscripts 1 and 2 respectively, and let us consider the process whereby some water molecules are inserted between the chains; then the partial free energy of the process of "filling" the junction points with a layer of water molecules and that of the corresponding rise in concentration of the salt solution as a result of water being given off will be:

Fig. 46. Diagrammatic picture of a junction point.

$$\triangle F_1 = \triangle H_1 - T\triangle S_1 \text{ and } \triangle F_2 = \triangle H_2 - T\triangle S_2$$

and hence, their sum, determining whether or not the process will spontaneously occur, is

$$\triangle F_{total} = \triangle F_1 - \triangle F_2 = \triangle H_1 - \triangle H_2 + T\triangle S_1 - T\triangle S_2 \tag{22}$$

Considering a series of salts, $T\triangle S_2$ is constant and $\triangle H_2$ becomes smaller according as the "lyotropic value" of the salt is greater[1]. Hence the chance that $+ \triangle F_{total}$ will be negative (as required if the process shall take place) is the greater the further the salt lies to the right in the lyotropic sequence[2].

As a matter of fact the action of neutral salts on gelatin and agar-agar[3] can be conceived, assuming that less and (or) looser junction points will be present according as the "lyotropic value" of the salt is higher.

In other words, we may say that the junction point spectrum of the gel (cf. p. 497) is changed in the corresponding sense. The idea suggested by HOFMEISTER, the

---

[1] It is recalled that $\triangle H_2$ represents the heat evolved in the process, with reversed sign.

[2] Observations of J. C. DERKSEN on the influence of salts on the side chain spacing of collagen, which is a measure of intramicellar swelling (penetration of water between the chains) are consistent with the picture set forth here (*Thesis*, Amsterdam 1935).

[3] Agar-agar must be also reckoned to the systems with electrolytic character as a result of the presence of sulphuric acid groups on the carbohydrate chains.

discoverer of the lyotropic series, that, owing to their own hydration, the ions withdraw water from the colloid, hence, touched the very heart of the question[1].

OUWELTJES [2], on account of his researches on the elastic extension and the liquefaction temperature of gelatin gels as affected by added salts, discusses at length the hypothesis that the partition between the amorphous and the crystalline portions of the gel framework is shifted according to the lyotropic value of the salt. The phenomena to be explained are also consistent with the assumption that there is a corresponding shift either of the number or of the cohesive energies of the non crystalline junction points in the amorphous portion of the gel, or of both. We shall revert to these problems later.

No matter whether OUWELTJES' assumption holds for gelatin gels, it would seem that it cannot be applied generally, since e. g. in cellulose gels, which are also subjected to "lyotropic" effects, it has been rather well established that variations of the proportion of crystalline substance (i. e. the fraction giving rise to a crystalline X-ray diffraction pattern) do not occur to a sensible, if to any, extent, either upon swelling and shrinking, or upon variation of the conditions of preparation [3].

The influence of salts on the solubility of macromolecular substances in water belongs to the same group of phenomena (salting-out effect) and is also subject to the HOFMEISTER rule. A typical example is cellulose xanthate, which is soluble in water and insoluble in salt solutions. If a film of liquid viscose, poured on a glas plate, is covered with a concentrated ammonium sulphate solution, the viscose solidifies to a cellulose xanthate gel. The degree of swelling of the gel depends on the concentration of the solution. It is the greater the more dilute the solution. If the solution is replaced by solutions of other salts of equal molar concentration, the degree of swelling varies according to the HOFMEISTER series. The changes are approximately reversible, if decomposition of the xanthate is prevented [4].

There seems to be little doubt that the solubility of cellulose in concentrated solutions of certain salts should be also explained on the basis of thermodynamic relationships similar to that symbolized by equation (22).

The conditions enhancing or reducing swelling which have been discussed here, finally resolve themselves into the equivalents of an increase or a reduction of "solubility", this being the same result as that arrived at in section a dealing with the systems without electrolyte character.

The concept solubility is used here in a somewhat different sense from the usual. We may perhaps rather express it thus: the spectrum of stable junction points between the molecular chains is shortened; the border line of solubility shifts towards the chain to chain contacts of greater energy.

b. 3. Some recent aspects of the mechanism of swelling

In the preceding section we have already met a theory on the mechanism of swelling of proteins in dependence of the hydrogen-ion concentration of the solution which may be termed the osmotic theory, and which seemed to be rather well supported

---

[1] Other authors have suggested that absorption of ions on the colloid plays a part. This would imply that the adsorption of the least hydrated ions is the most pronounced and thus finally enhances hydration of the colloid to a greater extent than strongly hydrated ions. This is, however, a mere assumption ad hoc. There are however certain indications that sometimes adsorption of ions may give rise to additional effects of salt action.

[2] J. L. OUWELTJES, Thesis, Amsterdam 1942.

[3] O. KRATKY and A. SEKORA, Kolloid-Z., 108 (1944), 169; P. H. HERMANS, Contribution to the physics of cellulose fibres, Amsterdam-New York 1946.

[4] Cf. H. G. BUNGENBERG DE JONG, Z. physik. Chem. (Cohen Festband) 1927, 205; P. H. HERMANS and A. J. DE LEEUW, Kolloid-Z., 81 (1937), 300.

by experimental data. In other instances of swelling the expansive tendency was connected with the swelling pressure. Both concepts, osmotic and swelling pressure, resolve themselves into the tendency of the system to reach greater dilution.

Very little has been explicitly said about the opposing cohesive tendency which counterbalances expansion and sets a limit to the swelling; neither have we entered into the question as to how the expansion of the framework structure of the gel should be visualized. In the last part of the preceding section the loosening of junction points by the agency of neutral salts has been connected with enhanced swelling. The question then arises as to why the gel expands, and as to whether the general diffusion tendency is the sole factor involved in this phenomenon.

In this connection a new trend of thought has recently emerged from progress in the physics of chain molecules (cf. Chapter IV by J. J. HERMANS). Although the development of these ideas is still in its infancy, it seems necessary now to refer briefly to them. The following exposition may convey some idea of the concepts involved.

According to the statistics of the chain molecule, the average distance r_0 between the ends of chains freely suspended in a liquid will be

$$r_0 = A \sqrt{N} \tag{23}$$

where A represents the length of the "statistical chain element", and N the number of chain elements per molecule. If the molecules were not convoluted and all the chain elements were aligned in one direction, their length l would be AN and hence

$$r_0 = \frac{l}{\sqrt{N}}$$

If by some agency the molecules are stretched beyond r_0, they will show a tendency to contract. Contrarily, if the distance between the ends of the chains is shortened, this will give rise to an expanding tendency. The chains behave like springs with an equilibrium length of r_0. Unlike ordinary springs, the recovery is due to an entropy effect and not to one of potential energy. In a sense, we might speak of the chains as being "entropy springs".

In gels having the structure of molecular networks the junctions points will strive apart on swelling and approach each other on shrinking. The length of the molecular chains interconnecting the junction points will vary accordingly and this will give rise to either a contracting or an expanding tendency depending up on the magnitude of the initial and final distance between the junction points as compared to r_0, the statistical equilibrium length which the chains would assume if they were "free". Using this principle, some considerable theoretical advance has recently been made in the field of rubberlike substances, where relatively very simple conditions are met with.

Moderately vulcanised rubberlike substances represent essentially molecular network structures. The junction points consist of chemical cross links of considerable and equal strength. They may be regarded as being fixed and permanent and not affected by the solvent or by temperature within a reasonable range. Moreover, the mutual cohesion between the hydrocarbon chains is very weak and can be neglected to a first approximation.

The geometrical pattern of the junction points is given in that their distribution may be considered to be random. The shape (degree of coiling) of the chains in the

original xerogel may be assumed to correspond to the most probable configuration of "free" chains.

The structural picture of the xerogel is that of a dense network structure consisting of tension-free randomly kinked chains, here and there interconnected by permanent junction points.

Upon swelling in a solvent, the network structure will expand, due to the tendency of the system as a whole to reach a state of greater dilution, that is a state of greater probability and entropy. The expansion, however, entails an extension of the molecular "entropy springs" between the junction points which, in its turn, entails a decrease in entropy of the individual chains.

Owing to the particular framework structure, the remarkable situation is met with that the entropy gain of dilution is counterbalanced by the entropy decrease bound up with the stretching of the molecular springs. This will set a limit to the swelling degree[1].

A successful quantitative treatment of this relatively simple picture, based upon statistical-mechanical derivations has been given by FLORY and REHNER[2] (see Chapter IV, p. 127). For the maximum degree of swelling q_m the equation

$$q_m^{5/3} = \frac{1 - 2\mu}{2} \cdot \frac{M_c}{V_1 \varrho} \tag{25}$$

was found, where V_1 is the molar volume of the solvent, ϱ the density of the undiluted rubber and M_c the average molecular weight of the chains between the junction points. The quantity μ is a constant characteristic for the system rubber-solvent and which takes care of the heat of mixing and other factors. Its value decreases according as more heat is evolved upon swelling[3].

Applied to a series of artificial rubbers (butyl rubber) with a different but known number of junction points, for which the average chain length between two junction points could, hence, be computed, the theory yielded excellent agreement between the maximum degree of swelling in a solvent calculated and that experimentally determined[4].

It should be noted, however, that limited swelling is by no means necessarily connected with the occurrence of a network structure with permanent crosslinks. Thermodynamical considerations of HUGGINS[5] and FLORY[6] have revealed that binary systems of a linear macromolecular component and a solvent may generally exhibit a region of demixture in two phases.

One phase then will consist of the pure solvent or very nearly so. This phase equilibrium is, however, extremely sensitive to temperature and variation of the solvent nature[7]. Upon even slight variations of one of the two factors, complete solution may occur. This is, of course, not the case in instances of limited swelling, bound up with the occurrence of a permanent network structure. Here, strong chemical

[1] J. FRENKEL, Acta Physicochim. USSR, 9 (1938), 235.
[2] P. J. FLORY and J. REHNER, J. Chem. Phys., 11 (1943), 512, 521; P. J. FLORY, Chem. Rev., 35 (1944), 51.
[3] Cf. M. L. HUGGINS, J. Am. Chem. Soc., 64 (1942), 1712.
[4] It should be remarked, however, that measurements of H. STAUDINGER on the swelling of polystyrene of various degree of cross linking (due to the addition of divinylbenzene) rather conform with an equation of the form $q^3_m = cM_c$ (Cf. F. H. MÜLLER, Kolloid Z., 95 (1941), 181).
[5] J. M. HUGGINS (loc. cit.).
[6] P. J. FLORY (loc. cit.).
[7] Cf. J. N. BRÖNSTED and K. VOLQUARTZ, Trans. Faraday Soc., 35 (1936), 576; 36 (1940), 619.

bonds would have to be overcome before complete solution could be reached.

It will be the aim of future research to further develop these very elucidating achievements in the mechanism of swelling and to attempt their extension to more complicated instances.

Emphasis should be laid on the very intimate relation apparantly existing between the phenomena of swelling and deformation in systems of this kind. Mechanical extension or compression involve an anisotropic deformation of the very same network structure, whereby changes in the distances between the junction points, bound up with stretching or contraction of the molecular chains extending between them, occur too. *Swelling may be compared with a threedimensional isotropic extension.*
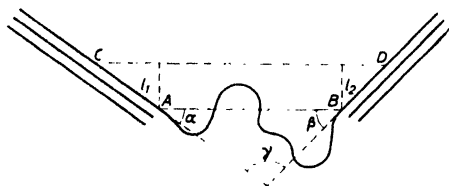
Fig. 47. Diagrammatic representation of two crystalline junction points connected by a common chain.

This intimate connection between swelling and deformation has also become apparent in recent work on the deformation of cellulose fibres (see Section 9c. 2, p. 638), where, however, the picture is a less simple one.

We shall now discuss some examples which go to show that also in theoretically less simple cases which cannot be treated quantitatively in an exact way, qualitative consideration of the same general principles may be of great help to reach a better understanding of many hitherto unexplained phenomena.

Let us consider the diagrammatic Fig. 47, representing two crystalline junction points connected by a molecular chain between the points A and B. (The figure and the principle of the reasoning are borrowed from OUWELTJES[1]. For simplicity, let us assume that the distance AB corresponds to equation (23) and hence, that the chain AB corresponds to the statistically most probable shape. If now, by a raise in temperature or by the addition of a salt, a partical "melting" of the crystallites occurs, whereby the sections AC and BD of a length l_1 and l_2 are freed from the lattice order and no longer bound to maintain their straight configuration, the total distance between the junction points will increase from AB to CD $=$ AB $+$ $l_1\cos \alpha + l_2\cos \beta$. If $l_1\cos \alpha + l_2\cos \beta = A \sqrt{N + N_1 + N_2} - \sqrt{N}$, where N_1 and N_2 are the number of chain elements corresponding to the liberated chainlengths l_1 and l_2, the system will again be in equilibrium. If, however, $l_1\cos \alpha + l_2\cos \beta$ is larger or smaller than $A \sqrt{N + N_1 + N_2} - \sqrt{N}$ there will be respectively an attraction or a repulsion between the junction points. The former alternative will be favoured by small values of α and β, hence by large values of the angle γ, and the latter will be favoured in the reverse case.

In isotropic gels, all values of γ between 0 and 180° will occur and a statistical mathematical treatment would be necessary to reveal as to which tendency predominates in a given case. Whichever the result will be, it is clear that always a new equilibrium condition is reached. The swelling, due to the tendency of the system as a whole to reach a state of greater dilution, will be set a limit at a certain degree of extension of the molecular entropy springs.

In gels with a marked uniaxial orientation of the crystallites the angles α and

---

[1] J. L. OUWELTJES, *Thesis*, Amsterdam 1942, p. 73.

β will be small and a contraction along the axis of orientation has to be expected whenever the percentage of amorphous fringes is enhanced at the cost of crystalline material. This now is exactly what is observed if orientated gels are allowed to swell, a phenomena termed swelling retraction[1].

Well known examples are the contraction of rayon fibres swollen in water upon submersion in concentrated solutions of sodium hydroxide and the contraction of crystallized raw rubber, stretched at a low temperature, upon swelling in a solvent like carbon bisulphide.

In the case just discussed, it was assumed that swelling is closely connected with a partial "melting" of the crystalline junction points. We can, of course, also imagine the case that the junction points remain unchanged. Let us consider a protein gel exposed to solutions with various pH. The expansive tendency will then vary according as the osmotic pressure changes (see p. 562). If the osmotic pressure increases, the gel will swell i.e. the average distance between the junction points increases. This will again give rise to a contractive tendency limiting the swelling.

Discussing the diagrammatical example of Fig. 47, we have assumed that, in the initial condition, the shape of the chains interconnecting the junction points of the gel corresponded to the statistically most probable one. Now there are obvious reasons to believe that this will only hold under exceptional conditions. It must be expected that the degree of convolution of the chains will largely depend upon the previous history of the gel. *We are confronted here with a very important point with regard to the structure and the behaviour of gels in general* and it is necessary now to outline briefly the fundamental ideas involved.

Let us consider a gel, say a gelatin or a cellulose gel, prepared by the gelatination of a solution. A framework consisting of junction points interlinked by molecular chains, having some unknown degree of convolution, extents itself throughout the gel. If now solvent is withdrawn from the gel, for instance by evaporation, the gel contracts. The junction points thereby come nearer to each other, which compels us to assume that the degree of convolution of the chains must increase. In any gel shrinking to some appreciable extent (and we know that very considerable shrinkage is a very common occurence), *some kind of folding or crumpling of the network structure is an inevitable consequence.*

The first to direct attention to this fact was BUSSE[2], and it was also stressed by KRUYT in a lecture on gels [3] where he stated: "Pendant la dessiccation, diminution du volume du gel, pliage des parois du réseau". The final structure of the contracted gel will largely depend upon the structure of the *primary gel* which was originally formed from the solution.

Moreover, the crumpled state of the framework implies that the shrunken gel has, so to speak, an "extra latent expansive tendency" of its own. In other words, upon addition of a suitable solvent, the gel is capable of swelling to a higher degree than another gel consisting of the same macromolecular substance which was originally formed at a smaller degree of dilution. In the former case the junction points of the gel will have to strive farther apart before equilibrium with the expansive tendency, due to swelling pressure, is reached.

[1] P. H. HERMANS and A. J. DE LEEUW, *Kolloid-Z.*, 81 (1937), 300; P. H. HERMANS, *Cellulosechemie*, 19 (1941), 117.

[2] W. F. BUSSE, *J. physic. Chem.*, 36 (1932), 2862.

[3] H. R. KRUYT, *Chim. et Ind.*, Vol. 42, No. 4 (1939).

These ideas also afford an explanation of some already long known phenomena, which have been designated as the *memorative faculty of gels*[1].

As early as 1927 GORTNER and HOFFMANN[2] prepared dry gelatin from gelatin gels, obtained by cooling gelatin solutions of various concentrations, and thereby found that, upon reswelling in water, the gelatin prepared from the more diluted primary gels absorbed more water than that prepared from more concentrated ones. The degree of swelling attained by the secondary gel tended to approach that of the corresponding primary gel. The same was observed by BUNGENBERG DE JONG and HENNEMANN in agar-agar gels[3]. It would seem as though in this cases the swelling of the dry gelatine and agar xerogels were nothing but a refilling of the original mesh work of the primary gel, the factor determining the degree of swelling being the tendency of the chains to recover the shape which they had in the primary gel.

Such a mechanism implies, however, that the energy of the new chain to chain contacts, which are doubtlessly formed as a result of the closer and closer packing of matter during the contraction and the final drying of the gels, remains below the energy of the upper end of the initial junction point spectrum. Swelling is then not impeded by the new junction points. The aforesaid condition is not fulfilled in cellulose gels, where the degree of swelling of the secundary gels, when swollen in water, appears to be independent of that of the primary gel. All cellulose gels reswell to approximately the same moderate degree[4]. The new junction points formed upon drying cannot be released by the mere action of water in this case. This is conceivable, since, unlike gelatin and agar-agar, cellulose is insoluble in water at any temperature between 0 and 100°. The primary aquo-gels are obtained by chemical decomposition of aqueous solutions of a cellulose derivative, like the cuprammonium compounds or the xanthate. The considerable degree of swelling which the gels exhibit, when prepared from solutions containing, say 4—8% cellulose, is to be regarded as being in a metastable one. Owing to the intrinsic insolubility of cellulose in water, a certain proportion of the new junction points formed during drying of the gels cannot be reopened by water. In Section 9c. 2 we shall see, however, that nevertheless many other properties of these gels largely depend on the degree of swelling of the primary gel, which is revealed by deformation experiments. The swollen secondary gels then behave as though the network of the primary gel were operative.

It would seem that this is due to the fact that a certain proportion of the additional junction points, formed upon drying, are torn loose during the deformation. The "memory" of gels for the original degree of swelling of the primary gel is particularly evident in those gels where the junction points are of the nature of primary valence bonds (cf. § 3d p. 494). After having been dried such gels reswell to exactly the original volume of the primary gel[5].

From the foregoing, it will be clear that the mechanism of swelling and contraction of gels in general involves various structural factors and will therefore often represent a rather complicated problem. Summarizing, these factors are
1. the degree of convolution (kinkyness) of the chains as compared to the statistically normal one,
2. the junction point spectrum,
3. the geometrical pattern of the junction points.

Upon changes of state, all these factors may vary simultaneously and influence each other. A satisfactory quantitative theoretical treatment will therefore be a very difficult task and, in attempts to tackle this problem, different lines of attack will have to be tried, according to the nature of the special case considered.

[1] P. H. HERMANS, *Kolloid-Z.*, 97 (1941), 231.
[2] R. A. GORTNER and W. F. HOFMANN, *J. Phys. Chem.*, 31 (1927), 464.
[3] H. G. BUNGENBERG DE JONG and J. P. HENNEMANN, *Kolloidchem. Beih.*, 36 (1932), 123.
[4] Unpublished experiments in the authors laboratory.
[5] See f. i. P. V. TAVEL, *Thesis*, Bern 1939; R. SIGNER and P. v. TAVEL, *Helv. chim. acta*, 26 (1943), 1972.

b. 4. Syneresis

Freshly prepared gels often show a tendency to contract; their volume decreases and solvent is spontaneously pressed out. This phenomenon has been termed syneresis by THOMAS GRAHAM. We have seen that, after the setting of a gel, equilibrium condition is usually not attained and further processes, generally embodied in the term "ageing", occur. Recrystallization and the formation of more junction points, or, in general terms, an extension of the junction point spectrum, may have a bearing on these phenomena. They are, however, of a rather complex nature and can by no means be considered as fully elucidated.

Syneresis is very common in gels. It is met with in macromolecular gels with water as well as in those with organic liquids as a solvent (e. g. rubber gels [1]). It is not confined to macromolecular systems but also occurs in gels formed from crystallizable substances, in the gels of the heavy metal hydroxides and the like [2].

The degree of syneresis often depends upon the concentration of the gel. Aqueous cellulose xanthate solutions, which are allowed to stand, set to a gel as a result of the spontaneous decomposition of the xanthate. Initially, the gel occupies the same volume as the original solution. If the experiment is carried out with solutions of various concentration, it is seen that beyond a certain dilution the gel contracts and the system, hence, separates in two phases, the gel and the liquid expelled from it. The phenomenon is the more pronounced the lower the concentration of the original solution. A similar behaviour was observed in silicic acid gels.

The influence of concentration, temperature and the addition of salts etc. on syneresis is, however, sometimes a rather complex one [3]. As a rule, added compounds which favour swelling diminish syneresis and vice versa [4].

It has also been found that the velocity of syneresis depends on the absolute dimensions of the gel fragment investigated. This is clear, if we take into consideration that the liquid which has to be pressed out must overcome a greater resistance in large fragments of the gel.

b. 5. Hysteresis phenomena in the preparation of xerogels

The hysteresis phenomena, occurring in the withdrawal of the solvent from a gel by evaporation, have been treated of in Sections a 3 β and γ (p. 529 and 536). Mention should still be made of a hysteresis phenomenon with respect to the final volume attained by the dry xerogel.

We have already seen that the astonishing capacity of the framework in diluted gels to undergo a very considerable degree of contraction is, yet, a limited one. Dried silicic acid gels retain a pore volume in the order of 60% (p. 537); cellulose gels are capable of reaching higher densities, but the densest possible packing is not reached (p. 538). Yet, it is obvious that the crumpling and folding up of the network structure (p. 571) must proceed to a very high degree in this case, and it is not surprising that there is a limit to this process. The degree to which the density of packing can be increased by withdrawal of the solvent will be bound up with the flexibility of the framework.

Gels like cellulose and gelatin become hard and brittle upon drying. (Completely dried cellulose can e. g. be hammered to a powder [5]). This can be interpreted thus. In the final phases of contraction the chains come so near to each other that the chance to form new chain to chain contacts increases enormously, and so many additonal junction points are formed that the mutual mobility of the chains is seriously hampered. Moreover, in cellulose and gelatin, the withdrawal of water increases the cohesion between the chains. The framework thus rendered rigid is devoid of the capacity to contract further.

[1] e. g. LE BLANC and KRÖGER, Z. Elektrochem., 27 (1921), 335.
[2] S. PRAHASH and N. R. DHAR, J. Ind. Chem. Soc., 7 (1930), 417.
[3] Cf. e. g. A. KUHN, Kolloid-Z., 46 (1928), 299.
[4] H. G. BUNGENBERG DE JONG, Kolloidchem. Beih., 29 (1929), 454.
[5] P. H. HERMANS, Cellulosechemie, 18 (1940), 97; also see pag. 627.

HUBERT, MATTHES and WEISBROD [1] have studied this phenomenon in cellulose and shown that the final degree of contraction can be increased if the drying is carried out repeatedly, or, even quicker, if special conditions are applied (drying at high temperatures and at precisely controlled humidity conditions). A higher final density of the xerogel can thus be reached [2], and such xerogels show a reduced swelling power when immersed in water. Whereas ordinary preparations of regenerated cellulose (rayon, cellophane) absorb 0.8 to 1 g of water per g cellulose, gels treated according to HUBERT, MATTHES and WEISBROD will only absorb down to 0.5 g. An interesting fact, reported by these authors, is that swelling of cellulose fibres in liquid ammonia at low temperatures, followed by evaporation of the latter, very quickly leads to the same result.

A similar effect could be reached in drying gelatin and casein gels. These effects, which are of some interest from a technical point of view, can be interpreted in that the junction point spectrum of the gel is broadened and more permanent junction points, resistant to the action of water, are being formed as a result of the special treatment. The phenomenon can be classified as representing a hysteresis of contraction.

That the effect is actually due to the formation of relatively few extra junction points, follows from the fact that, in cellulose, neither the X-ray diagram, nor the sorption isotherms of the treated samples show an appreciable change, indicating that the proportion of the crystalline and amorphous parts of the structure has not sensibly changed by the treatment. Only the limit of swelling is strongly affected. This is consistent with the point of view already put forward in the foregoing that the latter is determined by the particular structure of the framework.

b. 6. Volume contractions in swelling systems

It is a well known fact that, in the initial phases of swelling of a xerogel, the volume of the swollen gel is markedly less than the sum of the volume of the xerogel and that of the solvent absorbed. In the case of a gel like silica gel, admittedly containing a considerable amount of empty space, this is not surprising. In gels like cellulose and gelatin, where the occurrence of empty space, if present, is not so evident, the phenomenon has given rise to a great deal of discussion and experimental work. Since cellulose is again the best investigated case, we will take it as a representative example [3].

Cellulose was considered as a porous body. Determinations of its specific volume (reciprocal density) yielded different values when measured in different buoyant media. This was ascribed to the different power of various substances to penetrate into the pores, or to the different compression to which they were subjected as a result of adsorption. Helium gas, being an entirely indifferent substance with small molecules, was considered as being the most suitable medium to measure the specific volume of cellulose. Comparing the specific volumes measured in helium with those found in organic solvents like benzene on the one hand, and water on the other hand, the former were greater and the latter considerably smaller than the helium value. This was so interpreted that the organic solvents were not capable of completely penetrating into the pores, whereas water, as a result of its specific affinity to cellulose and the large attractive forces between the two substances, was compressed. This degree of compression was, however, so considerable that some authors could not agree with this explanation and other theories were put forward.

According to the present author, the subject should be treated from a different angle and can then be considerably simplified. The concept of density and specific volume are typically macroscopical ones, which loose their physical sense when

[1] E. HUBERT, A. MATTHES and K. WEISBROD, Kolloid-Z., 98 (1942), 173.
[2] This has been shown by the author. (Cf. the monograph cited below.
[3] For a detailed treatment of the subject and the literature concerned, see P. H. HERMANS, Contribution to the Physics of Cellulose Fibres, Amsterdam-New York 1946, p. 73.

applied to particles or pores falling within molecular dimensions. Now the so called porosity of xerogels, like cellulose, gelatin, agar-agar, doubtlessly falls within molecular dimensions. If a porosity of this kind should be measured with the aid of a medium consisting of elementary particles (molecules) of the same order of magnitude, the rules governing the packing of spheres of different dimensions should be applied rather than the macroscopical concepts of density and specific volume, which imply that the objects and the buoyant media concerned be structureless continua.

Cellulose and gelatin xerogels should be considered as a macromolecular glass or resin, comparable with other glasses, like for instance highly undercooled solid butanol, allowing for the presence of a crystalline portion in the former.

Amorphous substances show, as a rule, a lower density than the corresponding crystalline ones; the amorphous modification might, hence, in a sense, be considered as containing greater "pores". However, no more than one would suggest to measure the "porosity" of solid amorphous butanol by measuring its density in helium gas and comparing it with the density of the crystalline modification, should one endeavour to measure the "true density" of a xerogel in helium. If this were attempted, one would only measure the solubility of helium in the glass or in the gel. It would seem, however, that the gross density, which can be measured when using indifferent liquids not capable of penetrating into the gel, like benzene and carbon tetra-chloride in the case of cellulose, will very nearly approach the density in the proper sense[1].

The contractions observed in swelling have no other background than those frequently observed upon mixing of two liquids (like alcohol and water) or upon dissolution of a solvent in a liquid. Whatever may be the exact nature of these changes in volume, they can be understood from purely geometrical considerations, recalling that molecules of different size and shape are mixed. Similar contractions may for instance be also realized when mixing two piles of spheres of different radius.

Since swelling is intrinsically a process of dissolution, be it one of partial disso-lution, or even, in its initial phases bound up with the formation of addition compounds (or hydrates if water is concerned), the volume contractions do not represent any new or specific phenomenon[2]. In cellulose and gelatin gels, by far the main part of the total contraction coincides with the range of hydrate formation (cf. p. 541) and hence, may be connected with a chemical reaction, like the contraction occurring in the system sulphuric acid-water, or upon formation of crystalline hydrates of salts. It is known from numerous other examples that compound formation gives rise to a volume contraction.

The fact that the volume contraction of gelatin gels decreases upon raising the temperature at which it is measured[3], is consistent with the foregoing.

b. 7. Swelling and optical properties

α. *Refractive power*. In this section we shall exclusively deal with *isotropic* gels. The refraction of light in orientated anisotropic gels will be treated in section 8b (p. 585).

Since swelling is intrinsically nothing but a process of mixture, it should be

---

[1] As a matter of fact, the density of dry cellulose fibres, determined in such indifferent (non-swelling) liquids appears to be independent of the liquid chosen.

[2] It is interesting to note that a similar point of view was held by CONTONI about 80 years ago; *Rend. R. Inst. Lombardo*, 8 (1866) 135.

[3] THE SVEDBERG, *J. Am. Chem. Soc.*, 46 (1924), 2673.

expected that the refractive power of a swelling xerogel changes according to the same rules as those which have been found to govern the refractivity of homogeneous mixtures of two substances. This has been experimentally confirmed in the case of cellulose swelling in water.

Be g_1 and g_2 the weights, n_1 and n_2 the refractive indices and d_1 and d_2 the densities of the components of a mixture, further n_m and d_m the refractive index and the density of the mixture, then we may write

$$(g_1 + g_2) \frac{n_m - 1}{d_m} = g_1 \frac{n_1 - 1}{d_1} + g_2 \frac{n_2 - 1}{d_2} \tag{26}$$

This equation, which holds to a good approximation, represents nothing but the rule of GLADSTONE and DALE additively applied to the mixture[1]. For its verification the densities and the refractive indices of the components and those of the mixture, as well as the composition of the latter should be accurately measured.

Fig. 48. Refractive indices n_a of moist isotropic cellulose filaments against the water content a in g per g cellulose; (o) observed values, fully drawn curve calculated according to equation (26). (No attention should be paid to the dotted curve).

The equation accurately holds for mixtures of sulphuric acid and water over the whole range of composition, included those where the relative contraction is maximum. Measurements in the present authors laboratory on various specimen of isotropic cellulose[2] showed that it equally well applies to moist and swollen cellulose of any water content.

In Fig. 48 the observed refractive indices as a function of the water content are represented by dots, the curve is calculated with the aid of equation (26). (The dotted curve is an extrapolation of the final part of the experimental curve and cuts the ordinate at the apparent density of cellulose in water).

Applying the well known formula of LORENTZ-LORENZ one finds a less good agreement. This formula is based on a theory which does not account for the influence of the internal field. The GLADSTONE and DALE equation is a purely empirical relation which seems to cover the actual behaviour of systems of this kind very satisfactorily.

The theory of ONSAGER-BÖTTCHER is an attempt to improve the theory of the optical behaviour of mixtures, taking the influence of the internal field into consideration. Substituting the observed refractive indices in the somewhat more complicated equations of ONSAGER-BÖTTCHER[3] the average polarisability of the glucose residue and the radius r of the same could be calculated from the observations:

[1] Cf. C. DIETERICHI, *Ann. d. Physik*, 67 (1922) 337.
[2] P. H. HERMANS, *Contribution to the Physics of Cellulosefibres*, Amsterdam-New York 1946.
[3] Cf. C. J. F. BÖTTCHER, *Rec. trav. chim.*, 62 (1943) 329.

$$a = 11.4 \cdot 10^{-24}; r = 2.7 \cdot 10^{-8} \text{ cm}$$

According to the dimensions of the elementary cell of the cellulose crystal as derived from X-ray observation, the average radius of the glucose group should be 2.76 Å.

Results of this kind once more afford support to the view that the swelling of macromolecular systems in a solvent should be considered as a homogeneous mixture of the two components.

β *Light scattering*[1]. Various observations on the intensity of the TYNDALL cone in gels and its depolarisation have been reported, but few results were obtained throwing light on structural problems.

ARISZ[2] found that the intensity of the TYNDALL light increases during the gelatination of gelatin solutions. Similar observations were reported in agar-agar by KRISHNAMURTI[3]. HATSCHEK and HUMPHRY[4] found that an agar-agar solution prepared at high temperature and cooled to 62° was limpid in transmitted as well as in reflected light, whereas a gel of the same concentration heated to the same temperature is only somewhat opalescent in reflected light.

KRAEMER and DEXTER[5] found that the TYNDALL light of gelatin gels prepared from 1% solutions depends on the pH of the solution, showing a maximum at the iso-electric point. Moreover, different values were obtained when the temperature at which the gelatination was effected was varied. Especially in the isoelectric point, the intensity of the scattered light increased when the temperature of setting was lower.

It would seem that the variations observed are connected with the size of the crystallites formed, these being the discontinuities in the gel structure which may be of the proper size to give rise to light scattering.

b. 8. Swelling and X-ray diffraction pattern

The contributions of the examination with X-rays to the problems of gel structure and swelling have been particularly important. Various applications mentioned in the preceeding sections will serve to demonstrate this statement (p. 549) and few need be added at this place[6]. The most outstanding contributions were those concerning the question as to whether or not crystalline structures are present, under what conditions they appear or disappear, and what are the changes to which the lattice structure of the crystalline component is subjected under various conditions. Reliable quantitative information on the size and the shape of the crystallites which might, in theory, also be derived from X-ray diffraction patterns under favourable conditions, have been seldom, if ever, obtained.

Some interesting new aspects are provided by the application of the method of low angle scattering (cf. Vol. I, Chap. I) whose applications are however still in their infancy.

According to the theory developed by KRATKY[7] the intensity of low angle

[1] For the general theory of light scattering we refer to Part I, Chapter III.

[2] L. ARISZ, *Kolloidchem. Beih.*, 7 (1915), 1.

[3] K. KRISHNAMURTI, *Nature*, 124 (1929), 690.

[4] E. HATSCHEK and R. H. HUMPHRY, *Trans. Faraday Soc.*, 20 (1924), 18.

[5] E. O. KRAEMER and DEXTER, *J. Phys. Chem.*, 31 (1927), 764.

[6] We may refer once more to the monograph by J. R. KATZ, *Röntgenspektropgraphie als Untersuchungsmethode*, Berlin-Wien 1934.

[7] O. KRATKY, *Naturwiss.*, 26 (1938), 94.

scattering in a system consisting of closely packed rodled shaped crystallites will be very small, but will increase according as the average distance between the individual crystallites becomes greater. (This statement holds as long as the distance between the rodlets does not considerably exceed the lateral dimensions of the latter).

This is exactly what has been found for isotropic cellulose fibres of various degree of swelling [1]. The air-dry fibre shows very little or no small angle scattering at all, but the latter appears if the fibre is allowed to swell in water.

Fig. 49.   X-ray low angle scattering of cellulose filaments imbibed with a) ethyliodide, b) water

The intensity of the low angle scattering must also depend upon the absolute difference between the electronic densities of the crystallites and the medium between them [2]. For water this difference is 0.294, for toluene it is 0.379, for ethyliodide 0.016 and for methylene iodide 0.057. By imbibing cellulose filaments swollen in water first with alcohol and then by other organic liquids, KRATKY et al showed that the experiments are in qualitative agreement with these figures. As Fig. 49 shows, the intensity was almost zero in ethyliodide and very intense in water [3].

The method permits in principle the determination of the electronic density of the crystallites in a swelling system, in that the imbition liquid is sought yielding a minimum intensity. In turn, the electronic density of the crystallites may provide useful information on their chemical composition.

The foregoing applies exclusively if intermicellar swelling is concerned. In cases in intramicellar swelling, a considerable intensity of the low angle scattering is not to be expected. It was shown by KRATKY that in swelling cellulose acetate in various

[1] O. KRATKY, A. SEKORA, and R. TREER, Z. Elektrochem., 48 (1942), 587.

[2] The electronic density d is calculated from the equation $d_E = \frac{d}{M}\Sigma Z$, where d and M are the density and the molecular weight of the substance (in high polymers the molecular weight of the monomeric residue) and ΣZ the added atomic numbers of all the atoms contained in the molecule.

[3] These pictures were taken with a special camera designed for the purpose by KRATKY (Cf. Z. Elektrochem., 48 (1942), 409). The black line in the centre of the photographs is caused by the flat ribbon-shaped primary pencil of X-rays and should be disregarded. The low angle scattering is seen at both sides of white patches on the photograph. The ordinary crystalline cellulose interferences are not seen in the picture, since they fall outside the field of vision.

organic solvents giving rise to intramicellar swelling, the intensity of low angle scattering was weak and did not show a similar dependence of electronic density.

If anisometric e. g. rodlet-shaped crystallites are orientated, the low angle scattering will have the greatest intensity in the direction perpendicular to that of the largest diameter of the particles. The photograph in Fig. 50, borrowed from KRATKY, shows that this effect actually occurs. Extending the investigation to cellulose filaments showing biaxial orientation KRATKY could prove that the crystallites have the form of long flat lamellae rather than that of rodlets, a result which has also been deduced from other observations.

We shall also mention at this place some of KRATKY's preliminary experiments on the low angle scattering of concentrated solutions of macromolecular substances. Ordinary liquids and low molecular solutions will not yield a low angle scattering. Molecularly dispersed solutions of linear polymers may give rise to low angle scattering, which will be the more intense, the more low-distance-order (cf. p. 496) combined with associations, or swarm formation of molecular chains, occurs. Fig. 51 shows the results

Fig. 50. Low angle scattering in orientated swollen cellulose filaments (direction of the fibre axis vertical) (Borrowed from KRATKY).

obtained by KRATKY on solutions of cellulose nitrate in acetone. In the graph the intensity I is plotted against the distance r from the centre of the diagram. The time of exposure of the 6% solution has been taken twice as long as that of the 12% solution in order to bring both curves to a comparable scale. The maximum in the right hand part of the curves is due to a broad interference ring of the solvent. We can read from the diagram that the intensity of the curve for pure acetone diminishes towards the centre of the diagram (no low angle scattering) whereas that of the solutions increases considerably; mostly in the case of the 12% solution. According to KRATKY this would prove that effects like swarm formation and low distance order increase with the concentration of the solution.

In a similar way KRATKY showed that molecular aggregation increases, if cellulose xanthate swollen in sodium hydroxide is allowed to stand[1]. In the course of time the xanthate groups are split off and finally a cellulose gel is formed. (This process is termed in practice the ripening of viscose).

---

[1] The solution investigated contained 24.4% cellulose and was, hence, a very concentrated one, representing a gel rather than a solution.
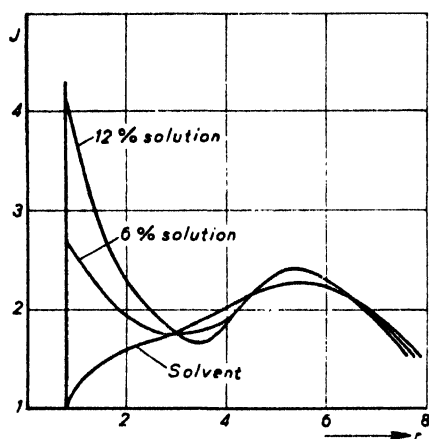
Fig. 51. Curves showing the observed intensity of low angle scattering in pure acetone and in 6 and 12% solutions of cellulose nitrate in this solvent (reduced on comparable scale).

STEURER[1] has shown by osmotic and other measurements that in solutions of ethyl cellulose in benzene association between the molecules occurs, and that this association can be broken by the addition of a small proportion of ethyl alcohol to the solvent. KRATKY could show that the addition of alcohol to the solution yielded a distinct lowering of the intensity of low angle scattering. Meanwhile it can be deduced from these investigations that even a limited association (in the case concerned it was found by STEURER that the osmotic activity of the solution was $1/2$ to $1/3$ of that corresponding to the actual number of molecules present) will give rise to measurable changes of low angle scattering intensity.

These examples may demonstrate the importance of the continuation of investigations in this field.

§ 7.   THE OPEN STRUCTURE OF GELS

In this section we shall briefly refer to some phenomena, which are connected with the typical micro-porosity of swollen gels and the continuity of the anastomizing spaces enclosed by the frame work structure of the gel-forming component.

a.   Diffusion and chemical reactions in gels as a matrix

The diffusion velocity of small molecules, soluble in the liquid component of swollen gels and not reacting in any way with the components of the gel, is almost equal to that in the pure liquid, provided the degree of swelling be not too small[2]. This is entirely consistent with the picture of gel structure developed. The same holds for the mobility of dissolved ions in an electric field. A great proportion of the gel volume is occupied by liquid in the normal condition and exhibiting normal properties. According as the degree of swelling decreases, the velocity of diffusion will of course diminish too, since a greater and greater relative part of the gel volume is occupied by the solid network structure[3].

A difference from diffusion in an ordinary liquid is, that diffusion in a gel is less liable to disturbance by mechanical influences or by convection currents. Several interesting applications have been made of this fact, for instance in the production of large crystals of inorganic substances[4]. Observations made on gels in this connection are of importance for the explanation of a number of mineralogical and geological

---

[1] E. STEURER, Z. physik. Chem., A 190 (1941) 1,16.
[2] H. BECHHOLD and A. ZIEGLER, Z. physik. Chem., 56 (1906), 105.
[3] We have already more than once pointed to the fact that, contrary to earlier popular conceptions, the amount of bound or immobilized solvent in gels is relatively small; cf. e. g. A. DOBRY, J. chim. phys., 35 (1938), 20; SIBREE, Trans. Faraday Soc., 26 (1930), 26; 27 (1931), 161.
[4] For a survey, see H. N. HOLMES in J. ALEXANDER's Colloid Chemistry, New York 1926.

problems. It needs not to be said that the study of diffusion in gels is also of great interest in biology.

If the diffusing substance reacts with the substance forming the gel frame, or if both are electrolytes, the phenomena become more complicated and this also applies if the molecules of the diffusing substance are not small as compared to the width of the channels of liquid in the gel. It will be clear that these phenomena interfere with the methods of "ultrafiltration". The membranes used for this purpose are typical gels.

Although a great many investigations on these and related subjects, dealing with the diffusion in gels and the permeability of membranes have been reported, the fundamental aspects of the subject can by no means be considered as fully elucidated. However important this field of research may be in more than one respect, it would seem that little additional fundamental information on the problems of gel structure may be derived from it. We shall therefore forbear from treating these items in this chapter and refer to the literature [1].

The same applies to several interesting phenomena observed if chemical reactions, giving rise to the formation of a precipitate, are allowed to take place in gels as a matrix [2].

b. Substitution of the liquid component by other solvents

A general property of swollen gels is that the liquid component may be easily interchanged by another liquid. If the second liquid is miscible with the liquid in the gel, the process will take place upon submersion of the gel in the former. If, on the other hand, water is to be replaced by a hydrocarbon, the aquo-gel may be first transformed into an alcohol gel and the alcohol replaced by the hydrocarbon. Such experiments were carried out by THOMAS GRAHAM with silicic acid and other gels [3], and later by O. BÜTSCHLI. If the second liquid is not a specific swelling agent for the gel, there is, generally, no or little change in volume of the gel at the end of the experiment. GRAHAM observed that silicic acid gels, when at once laid in alcohol, show some contraction. If, however, the aquo-gel is transferred into alcohol water solutions of gradually increasing concentration, no contraction is observed. The contraction observed in the first case is due to osmotic withdrawal of water from the aquo-gel by the alcohol.

Exactly the same phenomena are observed when freshly prepared highly swollen aquo-gels of cellulose are immersed in alcohol. If such a cellulose gel is laid in glycerol, a still greater contraction occurs, which can, however, be avoided if the gel is gradually treated with glycerol-water mixtures of increasing concentration [4]. If such precautions are taken, almost any gel can be "filled" with almost any liquid.

It is a very remarkable fact that a gel filled with a foreign liquid often shows

[1] See e. g. R. E. LIESEGANG, *Diffusion in jellies*, in J. ALEXANDER's *Colloid Chemistry*, New York 1926, p. 783; J. DUCLAUX, *Diffusion dans les gels et les solides*, Paris 1936. We further refer to the systematical studies of E. MANEGOLD on the structure and the permeability of membranes, *Kolloid-Z.* 80 (1937), 253 and numerous later papers in this journal in 1938 and 1939.

[2] Cf. e. g. S. C. BRADFORD in J. ALEXANDER's *Colloid Chemistry*, New York 1926, p. 790.

[3] TH. GRAHAM, *J. Chem. Soc. London*, 3517 (1864), 318; cf. E. HATSCHEK, *The foundations of colloid chemistry*, London 1925, p. 95.

[4] P. H. HERMANS, *Contribution to the Physics of Cellulose fibres*, Amsterdam-New York 1946.

a modified rigidity. It seems as though the stiffness of the framework has changed. Cellulose gels filled with alcohol or benzene are harder and less flexible than the original aquo-gels. If the foreign liquid is allowed to evaporate, eventually at high temperatures, the gels do not contract as easily as the aquo-gels did. Very often the gels then become opaque or even turn chalky white. The framework is not capable of folding up to the same extent as in the aquo-gels. In this way HERMANS and DE LEEUW prepared apparently "dry" cellulose gels, which contained only half as much cellulose per unit volume as dried aquo-gels.[1] Such gels retain, however, a considerable quantity of the organic solvent, which cannot be completely removed even after prolonged heating. The same authors described cellulose filaments, which, even after having been subjected to a temperature of 105° for many hours, still retained ethyl ether to more than 60% of their weight. Similar observations were reported by STAUDINGER and DÖHLE[2].

It would seem that steric factors play a part in these phenomena. Below a certain degree of swelling the molecular chains of the gel framework become very densely packed. Solvent molecules, which occur between them, may remain enclosed as in a cage with a grating too narrow to let them pass. Only such molecules will be able to pass from cage to cage whose affinity enables them occasionally to push their way through the grating by force of energetic interaction with the chains. This is the manner in which, for instance, water molecules can, though slowly, diffuse through cellulose gels with a low water content (cf. p. 545). Molecules which have no affinity for the substance of the framework, as e. g. benzene and ether in cellulose, when artificially introduced in a swollen gel, are unable to escape completely from it upon drying at high temperatures. It is a well known fact that the last traces of solvents are always very difficult to remove from gels.

Interesting experiments have been performed by KÜTLER[3]. He heated organogels of silicic acid in an autoclave beyond the critical temperature of the solvent and then allowed the latter to evaporate by releasing the pressure from the autoclave. He so obtained "dry" silicic acid gels which retained the volume of the original organogel and showed a very low spec. weight (e. g. 0.02). These objects were opaque and showed unusually good heat insulating power, since the capillary spaces contained in them were smaller than the wave length of heat radiation. Analogous results can be obtained with gelatin, agar-agar and cellulose gels, if organic liquids with a critical temperature lying below the decomposition temperature of the macromolecular component, like propane, butane and pentane are employed. The very voluminous xerogels so obtained represent, so to say, the original framework of the gel in which the liquid is replaced by air. Upon reswelling in water and drying once more, ordinary dense xerogels are again obtained.

A closer investigation of these phenomena might be of considerable interest. It is not yet quite clear why the gels, when subjected to the special procedure of "drying" referred to above, do not contract. Perhaps this is an indication that the theory discussed on p. 536, assuming that the capillary pull of the liquid menisci at the surface of the gel plays a part in the process of contraction, is correct. Beyond the critical temperature of the liquid these menisci no longer exist.

[1] P. H. HERMANS and A. J. DE LEEUW, Kolloid-Z., 82 (1938), 63.
[2] H. STAUDINGER and W. DÖHLE, J. prakt. Chem., 161 (1943), 219.
[3] KÜTLER, J. Phys. Chem., 36 (1932), 52.

c Chemical reactions of macromolecular systems in the gel-state

Chemical reactions in swollen gels whereby the macromolecular gel component is chemically changed will, of course, be of a topochemical character. The rate of the reaction will not only depend on the velocity coefficient of the reaction itself, but also on diffusion processes. The reagent has to diffuse into the gel to reach the scene of the reaction and reaction products will eventually have to leave that place by diffusion. Patches of the gel near its surface may react sooner than those farther away from the surface, if the reaction velocity is not very much smaller than the rate of diffusion.

The molecular chain sections lying more or less free in the amorphous portion of the gel will, as a rule, be affected sooner than those involved in the formation of junction points of some extension. Just as in swelling, it will depend up on the free energy change of the reaction as to how far the junction point spectrum is affected by the reaction, and the latter may stop at a certain point of the spectrum. In gels with distinctly crystalline junction points, which yield a crystalline X-ray diffraction pattern, examination of the latter will reveal whether or not the reaction also affects the crystallites. If the X-ray spectrum changes, we speak of an *intra*micellar reaction; if the reaction proceeds to a certain extent and the macromolecular component is partially transformed without a change in the X-ray photographs being observed, we speak of an *inter*micellar reaction.

In many cases where intramicellar reactions are possible it has been observed that first intermicellar reaction takes place, and intramicellar reaction only occurs after a longer lapse of time.

With the aid of these general principles many of the rather complicated phenomena, which have been observed in the chemical transformation of gels, may be satisfactorily interpreted.

As to the chemical transformation of xerogels, it will be clear that the course of the reaction will depend on whether or not the reagent is able to penetrate into the gel, in other words, on whether it is a swelling agent. If the gel does not swell in the reagent, or, if it only swells after the chemical transformation by the reagent is accomplished, the reaction will take a *macro-heterogeneous* course, gradually proceeding layer by layer from the gel surface towards deeper regions. An example of this is the acetylation of cellulose fibres by a solution of acetic anhydride in benzene. The process can be followed in the polarisation microscope, since the double refraction of triacetylcellulose is negative and that of cellulose positive, and the reaction is seen to proceed gradually from the surface to centre of the fibres.

An example of a transformation with a reagent which is a good swelling agent simultaneously, we may take the nitration of cellulose by a mixture of nitric and sulphuric acids. The whole fibre swells considerably and this swelling is even an intra micellar one, since an addition compound of cellulose and nitric acid (the so called KNECHT - compound) is formed, nitric acid molecules being taken up in the cellulose lattice and giving rise to another X-ray diffraction pattern. The subsequent nitration (substitution of hydroxyl groups by nitric acid) then proceeds homogeneously over the whole mass of the gel, and at the end of the reaction the nitrate groups are evenly distributed over the entire length of all chains. The final degree of substitution only depends up on, the composition of the acid mixture [1].

---

[1] Cf. e. g. F. D. MILES, *Trans. Faraday Soc.*, 29 (1939), 110; M. MATHIEU, *ibid.*, 29 (1939), 122.

In the case of a merely intermicellar reaction, even when taking place in a swollen gel, a substitution reaction will only reach certain sections of the chains and will leave other sections entirely unaffected. Several cases of this kind have been investigated, but they have often been otherwise and incorrectly interpreted[1]. Some very elucidating cases have been reported by CENTOLA[2] in cellulose reactions.

Acetylating cellulose fibres with acetic anhydride in which potassium acetate was dissolved, a reagent in which the fibres swell, CENTOLA found that a very considerable proportion of the total amount of hydroxyl groups could be esterified before the X-ray diffraction pattern of the cellulose was affected. When it changed at last, it became that of cellulose triacetate. The fibres then dissolved in chloroform, which is a solvent for triacetate. If the acetylation was stopped just before the X-ray diagram changed, the fibres could be swollen in chloroform but did not dissolve in this solvent. Only the amorphous fringes in the gel had been transformed into cellulose triacetate, but the junction points still consisted of unchanged cellulose. Though the gross-composition of the samples was near that of cellulose diacetate, they did not dissolve in the usual solvents for otherwise prepared ordinary cellulose diacetate, as obtained by partial desacetylation of cellulose triacetate in a homogeneous solution.

That actually triacetylated chain sections were present in the amorphous fringes of the partially acetylated samples was shown as follows. The samples were heated while under stress in vapour of boiling methanol. This caused the amorphous triacetate fringes of the gel to partially recrystallize, and an X-ray diffraction pattern was obtained revealing that a mixture of crystalline cellulose and crystalline triacetate was present. Reactions of this kind may be designated as *micro-heterogeneous* reactions.

Similarly different types of reactions have been also found in breakdown reactions whereby molecular chains are disrupted by chemical action. It will be clear that the mechanical strength of a gel, may be affected to a very different extent, when subjected to various breakdown reactions to the same gross extent, but taking a different course.

The phenomena referred to in this section are of considerable interest from a practical point of view in carrying out chemical transformations of macromolecular systems in the solid or the gel state, and in studying resistance against chemical breakdown[3].

§ 8. PHYSICAL PROPERTIES

CONNECTED WITH ORIENTATION IN GELS

a. General remarks

In practice, a great many anisotropic gels are met with. The anisotropy is invariably due to an orientation of the elements building up the framework structure. Anisotropy may reveal itself in various ways, as e. g. in the form of anisotropy of

---

[1] This applies for instance to the interpretations of K. HESS and cow. e. g. *Kolloid-Z.*, 68 (1934). 168; *Ergebn. Techn. Röntgenkunde*, 4 (1934), 21.

[2] G. CENTOLA, *Gazz. Chim. Ital.*, 65 (1935), 1015; *Atti X. Congr. intern. chim. Roma 1938*, Vol. IV, 123, 139, 722. Also see: D. VERMAAS and P. H. HERMANS, *J. Polymer Sci.* 2 (1947), 397, 406.

[3] A survey of the X-ray diffraction behaviour in cellulose reactions was given by W. A. SISSON, *Ind. Eng. Chem.*, 30 (1938), 530.

swelling, optical anisotropy (birefringence, dichroism) or mechanical anisotropy. The relative change of linear dimensions upon swelling or contraction, the refraction or absorption of light and the stress strain relations respectively, are then found to be different in different directions. If crystallites are present, orientation can be also read from the X-ray diffraction pattern. These phenomena of orientation are of outstanding importance from a scientific as well as from a practical point of view.

The majority of gel-like macromolecular systems occurring in the animal or vegetable kingdom is anisotropic. Orientation is due to the activity of the living matter from or by which they have been formed. The birefringence of the animal skin, cell walls, hairs and a great many other tissues has attracted the attention of investigators since the introduction of the polarizing microscope. Anisotropy of swelling, mechanical anisotropy and X-ray photographs indicating orientation have been frequently met with in such systems.

Anisotropic gels produced in the laboratory owe their orientation to a previous deformation. This may be either effected by mechanical deformation of an isotropic gel or by preventing the gel from isotropic contraction, for instance during drying. In both cases the random orientation of the structural elements of the frame work is changed into a more or less preferred one in one or two directions of space.

The systematic study of orientation phenomena has contributed a great deal to structural problems, and to our technical ability to produce products of a macromolecular nature exhibiting the properties desired for practical purposes.

b.   Optical effects of orientation

b. 1. Introduction

The optical phenomena bound up with orientation can conveniently be considered from the point of view of orientation of the molecules. Molecules are, as a rule, optically anisotropic particles, since their polarisability is different in various directions, except in a few cases. Systems consisting of oriented molecules, like crystals, will therefore be optically anisotropic, except regular crystals, which contain isotropic molecules isotropically arranged in space.

If isotropic molecules are anisotropically arranged in space, in that the intermolecular spacings are different in different directions, the system will exhibit optically anisotropy too. Such anisotropic arrangement may for instance ensue from a mechanical extension or compression in one direction or some other form of mechanical distortion. A regular crystal, when stretched in one direction, becomes birefringent.

Fig. 52.

This is the typical instance of "accidental birefringence". (cf. Vol. I, Chapter I, KRUYT.) The same applies to amorphous isotropic solids like common glass if subjected to mechanical deformation. This form of optical anisotropy is connected with relatively large tensions and very small deformations. Experience has shown that, in swollen gels, it may be taken for granted that no allowance needs to be made for this form
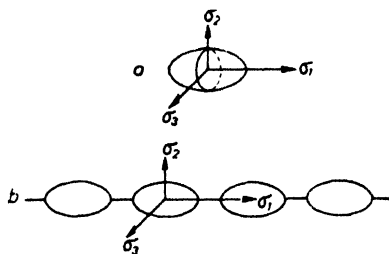
of optical anisotropy. In this case, tensions applied to the system may be always resolved into changes in molecular orientation.

Yet, other forms of anisotropic spatial arrangement of structural elements, which are themselves isotropic, may play an important part (Section 4, p. 594).

The optical properties of a molecule are conveniently expressed in terms of its polarizability in different directions, the polarizability being represented by a vector. The ends of the vectors will, generally, form the surface of a triaxial ellipsoid, which may be characterized by giving its three principal axes, being the "principal polarizabilities" of the molecule in these directions. (Fig. 52a). For a chain molecule (Fig.

Fig. 53. Monocrystal of a linear polymer.

52b) the principal polarisabilities of the monomeric residue should be given. The polarizability in the direction of the chain axis will be denoted by σ_1. If the molecules are orientated at random, the system will be isotropic, and its polarisability β_{iso} is represented by the equation

$$\beta_{iso} = N\sigma \tag{27}$$

where σ is the average polarisability of the molecule and N the number of molecules per unit volume. The refractive index n_{iso} of the system is then given by

$$\frac{n^2_{iso} - 1}{n^2_{iso} + 2} = \frac{4}{3}\pi\beta_{iso} \tag{28}$$

If the molecules have some preferred orientation, the polarizability of the system will be anisotropic and has to be calculated by vectorial addition of the anisotropies of the molecules, or that of the monomeric residues, when chain molecules are concerned. The anisotropy of a single chain molecule, having some arbitrary shape deviating from the straight configuration, must, of course, also be calculated by vectorial addition of the polarizabilities of the monomeric residues.

In a system consisting of perfectly orientated chain molecules, say a monocrystal as diagrammatically represented in Fig. 53, the three main polarisabilities of the crystal will be given by

$$\beta_1 = N_0\sigma_1 \qquad\qquad \beta_2 = N_0\sigma_2 \qquad\qquad \beta_3 = N_0\sigma_3 \tag{29}$$

where N_0 is the number of residues per cm³, and the three principal refractive indices of the crystal would be

$$\frac{n^2_1 - 1}{n^2_2 + 2} = \frac{4}{3}\pi\beta_1 \qquad \frac{n^2_2 - 1}{n^2_2 + 2} = \frac{4}{3}\pi\beta_2 \qquad \frac{n^2_3 - 1}{n^2_3 + 2} = \frac{4}{3}\pi\beta_3 \tag{30}$$

It should be noted, however, that these formulae imply the system to be a relatively diluted one. In denser systems (like in the case of a crystal) they will not necessarily hold, since molecules coming near to each other will mutually influence each others electrical behaviour. This influence of the so called "internal field" in denser systems can, as yet, not be cast into a general mathematical form. It would seem, that it is

negligible with respect to the *average* polarizability (equation (28)[1], but it must certainly be accounted for if the polarisabilities in different directions and, hence, if the birefringence of the system is concerned. From experiments in cellulose and rubber, a marked influence of the internal field can be deduced, and it interferes with the quantitative evaluation of orientation from optical data, as we shall see later.

The only adequate way to deal with orientation in macromolecular systems in a general quantitative manner remains, however, to express it in terms of the orientation of the structural units, the monomeric residues, and the anisotropy must then be expressed in terms of the polarizabilities of the latter. In each separate case a separate discussion will be necessary, in order to account for a possible influence of the internal field, which, as a rule, can not be calculated.

b. 2 *Various types of orientation.*

Two special types of orientation are of particular importance, uniaxial and biaxial orientation. In the former case, there is preferred orientation in one direction, in the latter case in two directions. The diagram in Fig. 54 may convey an idea of the two cases[2]. The structural units are represented by oblong lamellae. The length,
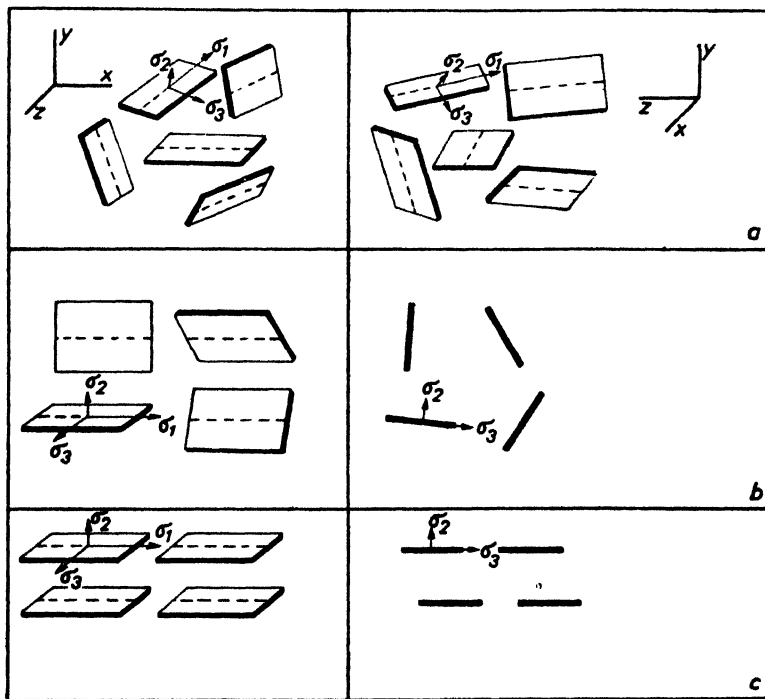
Fig. 54. Diagrammatic representation of a. random, b. uniaxial and c. biaxial orientation (in either case viewed in two directions perpendicular to each other, the direction of orientation included).

[1] In the case of cellulose at least, this follows from experiments of P. H. HERMANS and cow., cf. *Contribution to the Physics of Cellulose fibres*, Amsterdam-New York 1946.
[2] Cf. W. A. SISSON, *J. Phys. Chem.* 40 (1936) 343; 44 (1940), 513.

width and thickness of the platelets stand for the different polarizabilities in three directions. Fig. 54a shows *random orientation*. Viewed in two directions perpendicular to each other, the platelets are seen in a random spatial position. Fig. 54b shows the case of perfect *uniaxial orientation*, all the units being orientated with their axis of polarizability σ_1 pointing in one direction (x-axis). Viewed in the direction of orientation (Fig. 54b, right hand part) the axes of polarizability σ_2 and σ_3 are still orientated at random.

This kind of orientation is widespread in fibres. Continuous transitions between random orientation and perfect uniaxial orientation will exist according as the average orientation of the σ_1-axis of the particles is more and more a preferred one with respect to the X-axis. Fig. 54c shows perfect *biaxial orientation* of the particles. The σ_1, σ_2 and σ_3 axes of the particles now coincide with the X, Y and Z-axes of the coordinate system. Biaxial orientation frequently occurs in natural systems, e. g. in cell walls of plants and many other membrane-like objects. It is also widespread in artificial films like cellophane, rolled plastics etc.

The quantitative mathematical treatment of biaxial orientation is somewhat more complicated than that of uniaxial orientation and, entering into mathematical considerations later, we shall confine ourselves to the case of uniaxial orientation. It will be evident that systems with uniaxial and biaxial orientation will optically behave as uniaxial and biaxial crystals respectively.

Fig. 55. Uniaxial fibre orientation with prefered tangential orientation of the lamellae. With respect to the sector AMB, orientation is biaxial.

In systems with uniaxial orientation, the optical axis is coincident with the direction of principal orientation (X in Fig. 54b). The system will exhibit two polarizabilities $\beta_{,,}$ and $\beta_\perp$ and two principal refractive indices $n_{,,}$ and $n_\perp$, parallel and perpendicular to X respectively. The birefringence $\triangle$ is designated as being positive if $n_{,,} > n_\perp$ and negative if $n_{,,} < n_\perp$ and its magnitude is given by

$$\triangle = n_{,,} - n_\perp \tag{31}$$

Since the systems have symmetry of rotation around the X-axis, the relative magnitude of the elementary polarizabilities σ_2 and σ_3 is irrelevant with respect to the properties of the system and it is only their average value which counts. For the mathematical treatment we may therefore substitute an uniaxial elementary unit with the polarizability σ'_2 in all directions perpendicular to σ_1, where $\sigma'_2 = \frac{1}{2}(\sigma'_2 + \sigma_3)$.

A special form of uniaxial fibre orientation is diagrammatically shown in Fig. 55, where a cross section through the fibre is represented. The directions of the elementary polarizabilities σ_2 and σ_3 are orientated respectively radially and tangentially. The fibre as a whole can still be treated as a case of uniaxial orientation since σ_2 and σ_3 occur under all angles between 0° and 90° with respect to the Y and Z coordinates. If, however, a single sector like AMB would be cut from the fibre, this sector would have to be treated as biaxially orientated. This form of orientation occurs in many natural fibres. Considering fibres as a whole, there is no optical difference between the cases represented in Fig. 54b and Fig. 55. The elementary units in the diagram can then be substituted by rodlets and perfect orientation may be diagrammatically represented as in Fig. 56.

Imperfect uniaxial orientation is shown in Fig. 57 and, finally, another special case of uniaxial orientation is shown in Fig. 58. Here all the rodlets are at an angle of 90° to the fibre axis. This case is termed ringfibre structure. In fibres it does not occur, but it may occur in films (See Fig. 59). The rodlets then lie all parallel to the surface of the film (see cross-section in Fig. 59a). Viewed perpendicular to the film surface (Fig. 59b) there is random distribution. The optical axis is, of course, perpendicular to the film surface in this case. This case has been also designated as uniplanar orientation by W. A. SISSON (loc. cit., page 587). More frequently than rodlets, however, lamellae (ribbon shaped particles, like those shown in Fig. 54) occur. In that case the lamellar plane tends to be parallel to the film surface too, and then of course biaxial orientation enters the picture.

Fig. 56. Perfect orientation of rodlets.

Systems with biaxial orientation have three principal refractive indices in directions perpendicular to each other.

Fig. 57. Imperfect uniaxial orientation

In the foregoing we have considered various cases of orientation of individual units, either uniaxially or biaxially anisotropic ones. We have left open whether crystallites or monomeric residues were concerned. It will be clear that, in the latter case, some additional discussion is necessary, since monomeric residues are not individual particles free to orientate themselves wholly independently of each other, since they are linked together to form molecular chains. This subject has been treated in Chapter IV by J. J. HERMANS. The optical behaviour of chain molecules may be treated as though we were dealing with uniaxial rodlets having a length r (the vector connecting the two ends of the molecule) and an anisotropy of polarizability, given by the equation

Fig. 58. "Ringfibre" structure.

$$a_1 - a_2 = \frac{3}{5} \frac{r^2}{N A^2} (\sigma_1 - \sigma_2)$$

where σ_1 and σ_2 are the polarizabilities of the statistical chain elements considered as the elementary unit [1]. These imagined rodlets, hence, show a polarizability a_1 parallel, and a polarisability a_2 perpendicular to r, and may be substituted for the rodlets, shown in the preceding diagrams. A point of difference is, however, that, upon deformation of the system, the vectors r may not only change their position, but also their length r and, consequently, their anisotropy.

Fig. 59. Ringfibre structure in a film. It is the case designated as selective uniplanar orientation by SISSON.

It has already been stressed above, that orientation in man-made colloidal systems is always due to the previous agency of some external force applied to the system, from which a deformation of the system has arisen. If an isotropic gel is extended or compressed in one direction only, the resulting birefringence is uniaxial with the optical axis in the direction of the deformation. The sign of the birefringence depends on the kind of gel, but it is always opposite for extension and com-

---

[1] Cf. equation (4) on page 97.

pression. The same holds if a gel is hindered from swelling or contracting freely in one direction.

If deformation or inhibition of swelling (or contraction) is accomplished in two directions, the resulting anisotropy is, as a rule, biaxial, except in the case that the elementary particles themselves are uniaxially anisotropic, a case which will be very seldom, if ever, met with.

For cellulose, these phenomena have been comprehensively studied by W. A. SISSON (loc. cit., p. 587).

It is to be noted, that the effects of anisotropic deformation and anisotropic swelling run parallel. We shall see later, that and why the two factors are very intimately related in macromolecular systems (§ 9c. 2).

3. Imperfect uniaxial orientation and orientation factor.
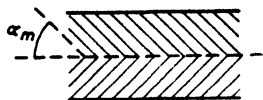
Fig. 60. Diagram of uniaxial orientation with uniform angle of orientation α_m.

If uniaxial orientation is imperfect (see Fig. 57), the birefringence of the system is a measure for its average orientation. We shall briefly discuss how this average orientation can be rationally expressed in one single figure, designated as the orientation factor f.

Let the value of this factor be 0 in the isotropic state, and 1 if the orientation is perfect. If the polarisabilities of the system parallel and perpendicular to the fibre axis (axis of preferred orientation) are $\beta_{,,}$ and $\beta_{\perp}$, and the polarisabilities at perfect orientation β_ε and $\beta\omega$, then the orientation factor is defined as follows:

$$f = \frac{\beta_{,,} - \beta_\perp}{\beta_\varepsilon - \beta\omega} \tag{32}$$

If $\beta_{,,} - \beta_\perp$ is small in comparison to $\beta_{,,}$ and $\beta_\perp$ (which is usually true) we may also write;

$$f = \frac{n_{,,} - n_\perp}{n_\varepsilon - n\omega} \tag{33}$$

thus expressing f in terms of refractive indices and birefringence. According to HERMANS and PLATZEK[1] the orientation factor may be connected with an average angle of orientation α_m of the elementary units of structure under consideration by the equation

$$f = 1 - \frac{3}{2} \sin^2 \alpha_m \tag{34}$$

At perfect orientation $f = 1$ and $\alpha_m = 0$.

The meaning of α_m can be explained thus. Let us imagine a case of imperfect uniaxial orientation of rodlets, thereby characterized that all the rodlets lie at the same angle α with respect to the optical axis of the system. (Fig. 60). Now, if we set the condition that the birefringence of this system shall be equal to that of the actual system under consideration, then the angle α shall be equal to α_m. For the

---

[1] P. H. HERMANS and P. PLATZEK, Kolloid, Z., 88 (1939), 68; cf. A. FREY-WYSSLING, Helv. chim. acta, 26 (1943) 833.

isotropic state $f = 0$ and a_m will be equal to arc sin $^2/_3 \cong 56°$. In the case of ring fibre structure (Fig. 48) $a_m = 90°$. The birefringence of the latter system will, of course, have an algebraic sign opposite of that of systems with a_m $\langle 56°$ [1]. The value of f is $-\frac{1}{2}$ in this case. In order to compute the orientation factor f from optical measurements, the birefringence of the system at perfect orientation $n_\varepsilon - n_\omega$ must, of course, be known. If one and the same object occurs in different degrees of uniaxial orientation, it is of interest to know the relationship between its main refractive indices and its orientation factor. This relationship, which can easily be derived from polarisability considerations [2] is given by the equations

$$n_{,,} = \tfrac{1}{3} \, [n_\varepsilon + 2n_\omega + 2f(n_\varepsilon - n_\omega)] \tag{35}$$

$$n_\perp = \tfrac{1}{3} \, [n_\varepsilon + 2n_\omega - f(n_\varepsilon - n_\omega)] \tag{36}$$

The refractive index of the isotropic object ($f = 0$) will, hence, be

$$n_{iso} = \tfrac{1}{3} \, (n_\varepsilon + 2n_\omega) = \tfrac{1}{3} \, (n_{,,} + 2n_\perp) \tag{37}$$

And equations (35) and (36) may therefore also be written in the form:

$$n_{,,} = n_{iso} + f(n_\varepsilon - n_{iso}) \tag{38}$$

$$n_\perp = n_{iso} - f(n_{iso} - n_\omega) \tag{39}$$

A graphical plot of $n_{,,}$ and $n_\perp$ as a function of f is shown in Fig. 61. It consists of two straight lines whose slopes are in the ratio of $2 : -1$.

These equations only hold under the condition that the density and the composition of the object be equal at all orientations. They then can be used to compute orientation from optical measurements.

Good agreement with experiments has been established in series of regenerated cellulose fibres from isotropic up to highly oriented ones, for which the values of n_ε and n_ω at ideal orientation were also computed with the aid of measurements on nearly perfectly oriented mercerized ramie fibres [3].

According to the well known rule of GLADSTONE and DALE (cf. p. 576), variations in density within these series could be accounted for by the equations

$$\frac{n_{iso} - 1}{d} = \text{constant} \tag{40}$$

$$\frac{n_{,,} - n_\perp}{d} = \text{constant at a given value of } f \tag{41}$$

Applied to cellulose fibres of different kind, it appeared, however, that equation (40) still held, but that equation (41) fails. The reason of this must be that the composition of the various fibres, though all consisting of cellulose, is not the same. Native fibres have a greater percentage of crystalline cellulose than regenerated ones (about 60%

---

[1] The general mathematical relation between the distribution function of orientation and f has been given by HERMANS and PLATZEK (loc. cit.).

[2] P. H. HERMANS, Contribution to the Physics of Cellulose fibres, Amsterdam-New York 1946 Other, more complicated formulae given by earlier authors, cited in this monograph, lead to the same numerical results and will therefore be omitted here.

[3] P. H. HERMANS, loc. cit.
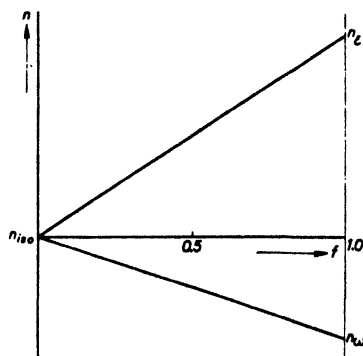
Fig. 61. Refractive indices of a body with uniaxial orientation, in dependence of the orientation factor f.

against 25%; cf. p. 541). Though the average polarizability and, hence, n_{iso} calculated from (37) remains independent of the percentage of crystalline substance and could be shown to be equal for all cellulose fibres, the anisotropy of polarizability and, hence, the double refraction is enhanced considerably according as the proportion of crystalline matter increases. The slope of the two curves shown in Fig. 61 becomes steeper, though maintaining the ratio — 2.

The refractive indices of native and regenerated cellulose at ideal orientation and in the isotropic condition (reduced to a density of 1.520) are collected in Table 8 (The percentage of crystalline matter is practically equal in all regenerated fibres).

TABLE 8

OPTICAL DATA OF DRY CELLULOSE FIBRES FOR $f = 1$ AND $f = 0$ AT A DENSITY OF 1.520

	$f = 1$			$f = 0$
	n_ε	n_ω	$n_\varepsilon - n_\omega$	n_{iso}
native ramie	1.590	1.519	0.071	1.543
regenerated fibres	1.575	1.525	0.050	1.542

In agreement with the foregoing, it could be shown experimentally that the double refraction of regenerated fibres changed if their crystalline portion (consisting of the crystalline modification cellulose II) was transformed into another modification (cellulose IV) without changing the orientation of tne fibre[1]. These effects, which impose certain restrictions as to the possibility of quantitative evaluation of orientation from optical measurements, arise from the influence of the internal field (cf. p. 586).

With the aid of equations (33) and (34) the orientation factor and the average angle of orientation of uniaxially oriented cellulose objects can be computed, if their birefringence is measured. This may be of some use in botanical work and in fibre research. It is necessary, however, to use the correct values of n_ε and n_ω. For practical purposes it will be sufficient to discriminate between native cellulose and regenerated cellulose. Further, the moisture content must be taken into account. The figures of n_ε and n_ω for the usual density of 1.555 in native and 1.520 in regenerated fibres, adjusted to the usual regain at 65 % rel. humidity which the fibres attain if first swollen in water and then conditioned in air of this humidity, are listed in Table 9.

An increase of the double refraction as a result of crystallisation has also been observed in rubber[2]. Quantitative theoretical considerations on the optical anisotropy

---

[1] This followed from parallel X-ray experiments.
[2] P. THIESSEN and W. WITTSTADT, Z. physik. Chem., B 29 (1925), 359; L. R. G. TRELOAR, Trans. Faraday Soc., 37 (1941), 84; 43 (1947), 284.

of rubber and other elastomers as a function of extension can, therefore, only be experimentally verified in the range of extensions, where crystallisation does not occur. Under conditions where crystallisation in rubber is abundant, as for instance

TABLE 9.

VALUES OF n_F AND n_ω OF CELLULOSE TO BE USED IN PRACTICAL MEASUREMENTS A1 65% REL. HUMIDITY

	n_F	n_ω	n_{iso}	$n_F - n_\omega$
native cellulose (d = 1.555) . . .	1.595	1.526⁵	1.549⁵	0.068⁵
regenerated cellulose (d = 1.520) .	1.554	1.512	1.526	0.042

at a temperature of 0°, the birefringence is a measure of the degree of crystallisation rather than of orientation, as follows from the work of TRELOAR (*loc. cit.*), the crystal lites being very well orientated even down to relatively small degrees of extension.

We have dealt rather broadly with these items, since, as yet, they have received but little attention.

In systems like cellulose fibres, extended rubber, polyamide fibres etc., the double refraction may be considered as being due to a superposition of the birefringence of the amorphous and that of the crystalline components.

Let the fractions of these components be $1 - x$ and x respectively, then the birefringence $\triangle$ of the system is represented by

$$\triangle = x\triangle_{cr} + (1 - x)\triangle_{am} \qquad (42)$$

where $\triangle_{cr}$ and $\triangle_{am}$ are the birefringences of the components. According to equation (33) we may further write

$$\triangle = xf_{cr}(\triangle_{cr})_i + (1 - x)f_{am}(\triangle_{am})_i \qquad (43)$$

where f_{cr} and f_{am} are the orientation factors of the crystalline and that of the amorphous component (which need not be equal[1], and $(\triangle_{cr})_i$ and $(\triangle_{am})_i$ represent the birefringence of the two components at ideal orientation (i. e. the value of $n_F - n_\omega$ in equation (33). The value of f_{cr} may be independently derived from X-ray experiments (see Section 624). In order to evaluate f_{am} separately, x, $(\triangle_i)_{cr}$ and $(\triangle_{am})_i$ should be known. We shall later meet an application of this equation [2].

It follows from the foregoing that the quantitative evaluation of orientation from optical data meets with various difficulties. It will only be possible in certain favourable cases, and with certain restrictions.

Still more complications are involved in the study of objects which are biaxially anisotropic, such as cell walls and artificial films. As an example of an optical study on films we may quote the work of J. SPENCE[3].

---

[1] In cellulose fibres it was shown, that a considerable difference between f_{cr} and f_{am} may occur (P. H. HERMANS, *loc. cit.*, p. 591). In rubber the difference is still much larger.

[2] It will be clear that the procedure of assigning a definite value of $(\triangle_{am})_i$ to the amorphous component is an approximation, since this quantity may also depend upon the density of packing and other factors.

[3] J. SPENCE, *J. Phys. Chem.*, 43 (1939), 865; 45 (1941), 401.

4. Structural birefringence in swollen gels.

The theory of a particular form of birefringence due to orientated structures of a special kind, we owe to WIENER[1]. A system of parallel rodlets or platelets embedded in a medium of different refractive power gives rise to a double refraction, even if the rodlets or platelets themselves are isotropic. The principal aspects of this phenomenon have already been treated in the general introduction (Vol. I, Chap. I) to which we may refer here. Instead of discriminating rodlet and platelet birefringence, it is preferable to simply take the sign of the bire-fringence as a criterion, since oblong lamellae, orientated with their longest axis in one direction, will also give rise to the same sign of birefringence as rodlets orientated in that direction (compare Fig. 62a and b), and *cet. par.* rodlets orientated
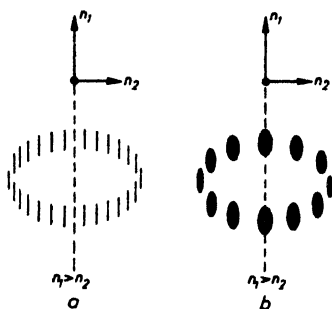
Fig. 62. Examples of *positive* structural birefringence $(n_1 > n_2)$
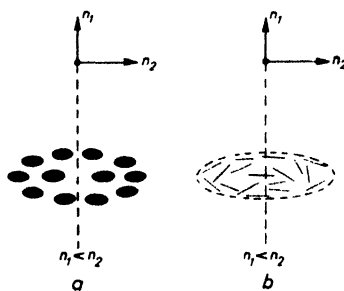
Fig. 63. Examples of *negative* structural birefringence $(n_1 < n_2)$

in parallel planes will yield the same type of birefringes as platelets orientated in parallel planes (Fig. 63a and b)[2]. We can better say that the direction of the greatest refractive power is parallel to the direction of orientation of the largest dimension of the particles. In several cases, where the occurrence of rodlet-shaped particles was concluded from structural double refraction, it was found later that, actually, orientated lamellae (Fig. 62b) were present.

Structural birefringence is detected by varying the refractive power of the imbibition liquid and measuring the double refraction in dependence of its refractive index. If a typical maximum or minimum curve is obtained, one may conclude that orientated anisodiametric particles occur[3]. A reversal of this rule is, however, not permitted, since the outcome of the experiment may largely depend upon the accessibility of the relevant structure to the various liquids used for the imbibition.

If the birefringence at the maximum or the minimum of the curves is markedly different from zero, it may be further concluded that the anisometric particles are birefringent themselves[4]. The remaining double refraction has been designated as the intrinsic birefringence of the particles.

[1] O. WIENER, *Abh. Sächs. Akad. Wiss.*, 32 (1912), 507; *Kolloidchem. Beih.*, 23 (1927), 189.

[2] H. AMBRONN and A. FREY, *Das Polarisationsmikroskop*, Leipzig 1926, p. 122.

[3] For an analysis of the shape of such WIENER curves cf. A. FREY-WYSSLING, *Kolloid-Z.*, 90 (1940), 33.

[4] For an analysis of the various possible combinations of particle shape and arrangement, see the book by AMBRONN and FREY cited previously, and W. J. SCHMIDT in E. ABDERHALDEN's, *Handbuch der biologischen Arbeitsmethoden*, Abt. 5 Teil 10 (1934), p. 435.

Structural birefringence has been met with in various typical orientated gels, e.g. in gels of gelatin[1], cellulose [2], cellulose nitrate [3], muscle myosin[4] and various objects of biological origin [5]. It demonstrates the presence of orientated polymolecular anisometric particles, since the theory of WIENER implies that definite phase boundaries shall exist between the particles and the surrounding medium. Orientated single chain molecules surrounded by a solvent will not give rise to a birefringence following the laws of WIENER. In as far as artificial gels are concerned, the relevant particles will, hence, be the crystalline junction points of the network structure and consist of a considerable number of orientated molecules. In agreement with this presumption, also an intrinsic birefringence was found in all these cases. (Structural birefringence in biological objects will also point to the presence of particles with much larger than molecular dimensions).

The question arises as to what is the contribution of the anisotropy of the amorphous fringe-like portion of the gel framework to the total birefringence. Since, in orientated gels, the amorphous parts will be also orientated, they certainly do add to the total double refraction, but, quantitatively, the question can not be answered at the present moment. The anisotropy of a molecular chain embedded in a solvent may be different according to the nature of the solvent.

A theory of the case was given bij SADRON [6]. An attempt to apply this theory to structural birefringence in gels was made by FREY-WYSSLING [7], who came to the conclusion that the structural component is *essentially negative* in the case of chain molecules. Since a satisfactory experimental verification of this theory has as yet not been achieved, we shall not enter into this subject. (The experiments of VERMAAS [8] discussed by FREY-WYSSLING in this connexion may be also interpreted in another way; cf. § 8b. 5).

When endeavouring to calculate the structural double refraction of a gel, another question is whether the refractive index of the matrix should be taken as being equal to that of the solvent, or to that of the mixture of the solvent and the amorphous portion of the gel. The polymolecular particles giving rise to the structural birefringence can not be considered as being imbedded in the pure solvent, but they are surrounded by a matrix consisting of a molecular dispersion of the amorphous fringes in the solvent. This may be one of the reasons why the theory of WIENER, when applied to gels, never leads to quantitatively correct results, but will only provide a qualitative or semi-qualitative guidance. Other reasons for this fact have already been mentioned in Vol. I, Chapter I.

Further, it should be noted that the theory of WIENER has been derived for diluted systems and that it can not be expected to hold in concentrated systems.

[1] H. AMBRONN, *Z. f. wiss. Mikr.*, 32 (1915), 43.

[2] H. AMBRONN, *Z. f. wiss. Mikr.*, 32 (1915), 56; *Kolloid-Z.*, 18 (1916), 278; A. MÖHRING, *Kolloidchem. Beih.*, 23 (1927), 162; P. H. HERMANS and P. PLATZEK, *Z. physik. Chem.*, A 185 (1939), 269; A. FREY-WYSSLING and H. SPEICH, *Helv. chim. acta*, 25 (1942), 1474.

[3] J. C. DERKSEN, J. R. KATZ, K. HESS and C. TROGUS, *Z. physik. Chem.*, A 149 (1930), 371.

[4] H. H. WEBER, *Pflügers Arch. ges. Physiol.*, 235 (1934), 205.

[5] e. g. DIEHL and G. VAN ITERSON, *Kolloid-Z.*, 73 (1935), 142; BEAR, SCHMITT and YOUNG, *Proc. Roy. Soc. London*, 833 (1937), 505; F. O. SCHMITT, *J. Appl. Phys.* 9 (1938), 109.

[6] G. J. SADRON, *J. Physique et Radium*, 8 (1937), 481.

[7] A. FREY-WYSSLING, *Helv. phys. acta*, 16 (1943), 437.

[8] D. VERMAAS, *Z. physik. Chem.*, B 52 (1942) 131.

According to the theory of WIENER (Cf. Vol. I, Chapter I and III) the refractive indices of a composite body consisting of parallel rodlets are given by

$$n_{,,}^{2} = \frac{n_1^{2} + (q-1)n_2^{2}}{q}; \quad n_{,}^{2} = n_2^{2}\frac{(q+1)n_1^{2} + (q-1)n_2^{2}}{(q+1)n_2^{2} + (q-1)n_1^{2}} \tag{45}$$

where n_1 and n_2 are the refractive indices of the rodlets and of the matrix respectively, and q is the degree of swelling[1]. If anisotropic rodlets are considered, n_1 may be taken equal to the average value of the two principal refractive indices of the rodlets[2].

In Fig. 64 the birefringence $n_{,,} - n_{\perp}$, calculated for certain values of n_1 and n_2 according to the equations (45), is plotted against the degree of swelling q.
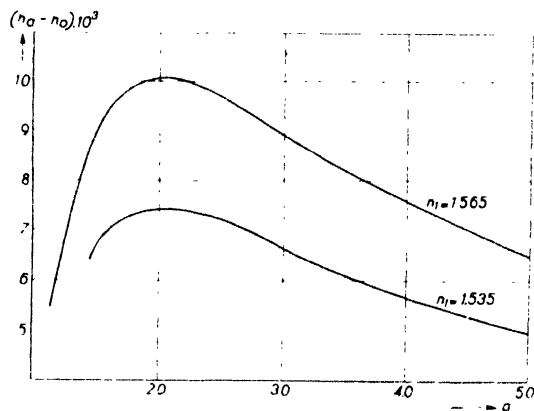
Fig. 64. WIENER component of birefringence in dependence on the degree of swelling q, calculated for the values of n_1 indicated and for $n_2 = 1.3333$.

It will be seen that the birefringence is maximum for $q = 2$, i. e. when the partial volumes of the particles and of the medium are equal. It is open to doubt, however, if the degree of dilution of the system at $q = 2$ is large enough to ensure correct applicability of the theory. However this may be, for values considerably below $q = 2$ departures will, at any rate, be possible. These facts have as yet not been stressed, and have already given rise to failures in the interpretation of the optical phenomena bound up with the initial phases of swelling[3].

In one case, the swelling of cellulose fibres in water, some more light has been thrown on the optical behaviour of the system by investigations of the present author and his collaborators[4], which we shall now briefly describe. It has already been mentioned in Section 6b. 7 (p. 575) that the change in refractive power of isotropic cellulose upon swelling in water follows the simple rule of additivity of the refractions $(n-1)/d$ of the components. It was established that, up to a water content of about 20 per cent, the same rule may be applied to the principal refractive indices of anisotropic fibres separately. This implies that the double refraction $n_{,,} - n_{\perp}$ of the fibre is simply inversely proportional to the degree of swelling p:

$$(n_{,,} - n_{\perp})_a = \frac{(n_{,,} - n_{\perp})_o}{q} \tag{46}$$

---

[1] The partial, volumes δ_1 and δ_2 of the components occurring in WIENER's equations can be thus expressed in terms of q:

$$\delta_1 = \frac{1}{q} \quad \text{and} \quad \delta_2 = \frac{q-1}{q}.$$

[2] A. MÖHRING, Kolloidchem. Beih., 23 (1927), 126.
[3] e. g. M. MEYER and A. FREY-WYSSLING, Helv. chim. acta, 18 (1935), 1428; P. H. HERMANS and P. PLATZEK, Rec. trav. chim., 58 (1937), 1001 who did not succeed in accounting for the optical behaviour of cellulose upon absorption of water.
[4] P. H. HERMANS, Contribution to the Physics of Cellulose fibres, Amsterdam 1946, p. 115.

where the subscripts a and o refer to a water content of a g per g of cellulose and to the dry fibre respectively. If more water than 0.2 g per g of cellulose is taken up ($q > 1.25$) the birefringence decreases somewhat less rapidly than indicated by equation (46) and consequently an additional positive component of double refraction, which might be due to a WIENER effect, reveals itself. The total birefringence at a degree of swelling q may then be written

$$(n_{//} - n_{\perp})_q = \frac{(n_{//} - n_{\perp})_o}{q} + (n_{//} - n_{\perp})_W \qquad (47)$$

where the last term stands for the WIENER-component of the double refraction, which, in theory, should be given by a calculation of $(n_{//})_W$ and $(n_{\perp})_W$ with the aid of the equation (45).

   Fig. 65 shows the course of the values of $(n_{//})_q$ and $(n_{\perp})_q$ in dependence on the water content a of the fibre, as found experimentally (dots), as calculated according to the additive rule of GLADSTONE and DALE (fully drawn lines), and as calculated according to equation (45) combined with equation (47) (broken line). The WIENER component of double refraction calculated according to (45), substituting for n_2 the refractive index of water, amounts to 0.0058 at 20% moisture content, whereas the observed value was only 0.002 (almost falling within experimental error). At $q = 2.12$, where the WIENER component should be almost at its maximum value, the observed WIENER component, derived from the departure from equation (46), was 0.0041, whereas 0.0085 is required by theory. Hence, a definite WIENER effect reveals itself here, but it remains smaller than half the theoretical value[1].

Fig. 65. $n_{//}$ and $n_{\perp}$ as a function of water content for highly oriented cellulose filaments. Broken curves represent the theory of WIENER, fully drawn curves equation (26), and dots the values observed.

   The present author has offered the following explanantion of the discrepancy: in the dry fibre phase boundaries may be supposed to exist between the crystalline regions and the amorphous ones. The difference in refractive power of the two phases then is, however, very small. The WIENER effect, being dependent upon the difference in refractive power of the crystalline particles and the surrounding medium, is too small as to be observable.

   In the moist fibre, the WIENER effect is not due to the difference in refractive power between the crystallites and the water, but to that between the former and the homogeneous mixture of cellulose and water composing the amorphous part of the swollen fibre. Since the decrease of the refractive power of cellulose in dependence of water absorption is very slow at first (cf. Fig. 48, p. 576), the WIENER effect will not become observable until a relatively high water content is reached,

----

   [1] According to unpublished calculations of J. J. HERMANS, WIENER's calculations. moreover, contain an error, and the second of the equations (45) should be

$$n_{\perp}{}^2 = n_2{}^2 + (1 + \frac{2}{q}) \frac{n_1{}^2 - n_2{}^2}{n_1{}^2 + n_2{}^2}.$$

The birefringence following from this corrected equation is, however, still greater than that according to equation (45).

and even then it will always remain smaller than that calculated by substituting the refractive indices of dry cellulose and water in equations (45). The "medium" still consists of a mixture of amorphous cellulose and water, having a much higher refractive power than pure water.

It may be expected that the optical behaviour of other macromolecular systems at swelling will be analogous to that of the system cellulose-water.

In the foregoing it was implicitly taken for granted that the orientation of the rodlet was perfect. (In the cellulose fibres treated of above the orientation factor was high enough to approximate this case). If there is only partial orientation, the WIENER-component of birefringence will still exist, but it will be accordingly smaller. A theory of the quantitative relationship with orientation, however, remains to be developed. In several experimental investigations it has been found that the WIENER component Δ_W as well as the component of intrinsic birefringe. Δ_I both run parallel with orientation[1]. Fig. 66, borrowed from experiments of D. VERMAAS, shows an example referring to filaments of cellulose nitrate, which were imbibed in alcoholic solutions of potassium mercuric iodide of increasing concentration and refractive power.

Fig. 66. WIENER imbibition curves of cellulose nitrate filaments in alcoholic potassium mercuric iodide solutions of increasing refractive index n_2 (Figures next to curves refer to degree of previous extension and run parallel with orientation.)

The intrinsic double refraction of cellulose nitrate being negative, the curves exhibit a maximum. It is seen that both the birefringence in the maximum and the WIENER component increase with orientation. KRATKY and PLATZEK[2], as well as the other authors just cited, have endeavoured to interpret the results of their measurements using the equation

$$\Delta_t = \frac{f \Delta_I}{q} + f \Delta_W \tag{48}$$

where Δ_t is the total birefringence and f the orientation factor, which was assumed to be equal for both components (compare also equation (47)). The correctness of this equation however, is still open to doubt, and so are some of the further conclusions arrived at in its application.

Finally, it is necessary to mention the fact that, in a given gel, regular WIENER curves are usually only obtained with series of certain selected imbibition liquids, while other liquids do not yield consistent results. Cellulose gels, for instance, can be imbibed with any organic solvent if the objects are first swollen in water, then laid in alcohol to replace the water and finally submerged in another organic liquid. However, regular WIENER curves are only obtained in alcohols and organic bases like toluidine and quinoline. Liquids like heptane, chloroform, ethyl iodide, carbon bisulphide, either

[1] P. H. HERMANS and P. PLATZEK, Z. physik. Chem., A 185 (1939), 269; H. R. KRUYT, D. VERMAAS and P. H. HERMANS, Kolloid Z., 100 (1942), 111; D. VERMAAS, Thesis, Utrecht 1941.
[2] O. KRATKY and P. PLATZEK, Kolloid-Z., 84 (1938) 268; 88 (1939) 98.

do not influence the birefringence at all, or very little [1]. It would seem that only liquids with a certain "affinity" for cellulose (containing polar groups) give rise to a normal WIENER-effect. The present author has offered an explanation for this fact elsewhere [2].

To conclude, many pitfalls may be met with when using the phenomena of structural birefringence in order to elucidate the structure of gels. Another case of that kind will be treated of in the next section.

5. Birefringence owing to orientation of solvent molecules.

Another factor, which may interfere with the birefringence of orientated swollen macromolecular systems, and which has, thus far, almost entirely been overlooked in all previous treatises on the subject, is bound up with the orientation of the molecules of the imbibition liquid. In a previous section, we have seen that a part of the molecules of a swelling agent may be rather tightly bound on the surface of the chains of the macromolecular substance and they will then, obviously, also occupy a fixed position relative to the chain axes. If the latter are orientated, the solvent molecules will show preferred orientation too and hence, may contribute to the total birefringence of the system. Attention was first directed to this possibility by HERMANS and PLATZEK (loc. cit.) and they deemed it probable that certain hitherto unexplained observations should be interpreted on this basis. MÖHRING [3] e. g. reported that the positive intrinsic birefringence of orientated gelatin gels changed sign in cresol-gelatin. Some very striking examples of this kind of birefringence were later found by D. VERMAAS [4], who investigated the birefringence of cellulose nitrate swollen in a great many organic liquids. In order to explain the total birefringence observed, he found himself compelled to add still another term to equation (48) leading to the expression:

$$\Delta_t = \frac{f \cdot \Delta_I}{q} + f \cdot \Delta_W + f \cdot \Delta_A \tag{49}$$

where Δ_A stands for the birefringence owing to orientation of the solvent, which he termed "adsorption birefringence" (Adsorptions-Doppelbrechung).

In certain liquids (especially those containing an aromatic nucleus) the term $f \Delta_A$ by far surpassed the two other ones and even sometimes caused a change of sign of the total birefringence! The intrinsic birefringence of the cellulose nitrate used was negative and, upon permeation with alcohols of increasing refractive power, or with alcoholic solutions of potassium mercuric iodide, normal WIENER curves were observed, the sign of the double refraction remaining negative over the entire range examined (cf. Fig. 66). Other liquids showed a tremendous departure from this behaviour. Some examples are shown in Table 10.

In the second and third columns the refractive index n of the liquids and the swelling degree q are listed. The total birefringence Δ_t observed is given in the fourth column. The intrinsic double refraction $f \Delta_I$ of the filaments was — 32.2 . 10^{-4} . In

[1] P. H. HERMANS and P. PLATZEK, Z. physik Chem., A 185 (1939) 269; A FREY-WYSSLING and H. SPEICH, Helv. chim. acta, 25 (1942) 1474.

[2] P. H. HERMANS, Contribution to the Physics of Cellulose fibres, Amsterdam 1946, p. 126.

[3] A. MÖHRING, Kolloidchem. Beih., 23 (1927), 152; cf. A. KÜNTZEL and F. PRAKKE, Biochem. Z. 266 (1933), 243.

[4] D. VERMAAS, Thesis, Utrecht 1941; Z. physik. Chem., B 52 (1942), 131; D. VERMAAS et al, Kolloid-Z., 100 (1942), 111.

the fifth and sixth column the values calculated for $f \cdot \varDelta_I / q$ and for the double refraction according to WIENER $f \cdot \varDelta_W$. Their sum should be equal to $\varDelta_t$ if equation (48) holds. We see that nothing of the kind is true. This is why VERMAAS added the term $f \cdot \varDelta_A$

TABLE 10

Solvent	n_z	q	$\varDelta_t$	$\dfrac{f \cdot \varDelta_I}{q}$	$f \cdot \varDelta_W$	$f \cdot \varDelta_A$
Trichloro-ethene . .	1.475	1.53	—20.9	—21.1	+2.1	—1.9
meta xylene	1.495	1.79	+21.9	—18.0	+0.8	+39.1
tetrachloro-ethane .	1.497	1.77	—41.2	—18.2	+0.8	+23.8
benzene	1.501	1.79	+25.7	—18.0	+0.3	+43.4
chlorobenzene . . .	1.525	1.79	+23.0	—18.0	0	+41.0
dibromo-ethane . .	1.538	1.75	—38.0	—18.4	+0.3	—19.9
benzylalcohol	1.540	1.75	+19.1	—18.4	+0.4	+37.1

in equation (49), which according to the last column of the table varies between — 20 · 10⁻⁴ and + 43.4 · 10⁻⁴ in the cases investigated! It is largest for the aromatic liquids. According to VERMAAS, in these liquids the molecules next to the cellulose nitrate chains would be orientated with the plane of the benzene nucleus parallel to the axis of the chain, thus giving rise to a reversal of the sign of the double refraction of the system[1].

These experiments show that *one will always have to bear in mind the possibility of a component of birefringence due to adsorption, when dealing with swollen gels*, a fact which sofar has been frequently overlooked. Particularly if this component is small, it will, however, not always be easily detectable. It may then be confounded with other effects and lead to erroneous conclusions.

It is a very striking fact that water, when taken up in cellulose (and apparantly also in other objects like gelatin), does not give rise to any detectable component of this kind, though some of the water molecules are certainly very tightly bound and are anisotropic themselves. This fact remains to be explained.

6. *Determination of birefringence*

For the details of optical measurements, we may refer to other textbooks[2]. We shall confine ourselves to recalling briefly that two general methods are used:

Firstly, one can measure the principal refractive indices of the object separately, using the immersion method of BECKE or that of SCHRÖDER VAN DER KOLK. In the case of uniaxial orientation the birefringence is equal to the difference of the two principal refractive indices.

The other method consists of measuring the phase difference γ in the polarizing microscope with the aid of a suitable compensating device. The birefringence is then given by

$$n_{,,} - d_{\perp} = \frac{\gamma \lambda}{d} \tag{49a}$$

[1] It should be remarked that the liquids concerned only show *intermicellar* swelling and no trace of *intramicellar* swelling. Hence, the term adsorption birefringence seems to be justified.

[2] *e. g.* H. AMBRONN and A. FREY, *Das Polarisationsmikroskop*, Leipzig 1926; F. RINNE und M. BEREK, *Anleitung zu optischen Untersuchungen mit dem Polarisationsmikroskop*, Leipzig 1934.

where λ is the wavelength of the light used and d the distance which the beam of light has travelled through the object expressed in the same unit of length[1].

7. *Dichroism*

Optical anisotropy in an object may also result in its having different sorptive power in different directions for polarised light. Such objects show the phenomenon designated as dichroism. Viewed under the microscope in monochromatic polarised light, the intensity of the transmitted light varies, if the orientation of plane of polarisation of the incident beam is changed. In white light a change in colour is observed.

Dichroism is to be expected, when dealing with orientated molecules of a coloured substance. It was *e. g.* observed by AMBRONN[2] in cellulose fibres coloured with certain dyestuffs. The quantitative treatment of dichroism is analogous to that of double refraction, substituting absorption coefficients k for refractive indices or polarizabilities.

Considering the case of uniaxial orientation, and assuming that the LAMBERT-BEER law is applicable, the intensities of the transmitted light with the plane of polarisation parallel and perpendicular to the optical axis will be:

$$I_{//} = I_o e^{-k_{//}d/\lambda} \qquad I_{\perp} = I_o e^{-k_{\perp}d\lambda} \tag{50}$$

where I_o is the intensity of the incident light, d the way the light has travelled through the object and λ the wavelength. Hence:

$$k_{//} - k_{\perp} = \lambda \frac{\ln I_{\perp}}{\ln I_{//}} \tag{51}$$

This quantity is a measure of the dichroism and can be computed from a determination of $I_{\perp}$ and $I_{//}$. The equation is analogous to equation (49a) in the preceeding section.

If I_o is also known, we may define another constant, which is independent of the wave length:

$$\frac{\ln I_{//}/I_o}{\ln I_{\perp}/I_o} = \frac{k_{//}}{k_{\perp}} - p \tag{52}$$

termed dichroitic constant by PRESTON[3]. The dichroism is largest for light of a wavelength near the absorption bands in the spectrum of the substance concerned.

Besides *intrinsic* dichroism, *structural* dichroism may occur too. The latter will be *e. g.* observed if coloured isotropic parallelized rodlets are imbedded in a matrix with a different light absorption. The theory has been treated also by WIENER[4]. The results are much more complicated than in the case of structural double refraction of colourless objects. The birefringence itself is also affected if one of the components of the composite body is coloured, and will for instance in general not be zero at equal refractive power of rodlets and matrix. We shall not enter into further details here.

Examples of structural dichroism were described by FREY-WYSSLING in fibres, dyed with colloidal metals like mercury, silver and gold[5]. Sometimes the metal particles lie in rows directed parallel to the fibre axis and thus give rise to structural dichroism.

[1] A trick facilitating such measurements in the case of relatively thick filaments was described by P. H. HERMANS and P. PLATZEK, *Z. physik. Chem.*, A 185 (1939), 260.

[2] H. AMBRONN, *Ber. deutsche Bot. Ges.*, 6 (1888), 25, 225.

[3] J. M. PRESTON, *J. Soc. Dyers and Colourists*, 47 (1931), 312.

[4] O. WIENER, *Kolloidchem. Beih.*, 23 (1927), 189.

[5] A. FREY, *Kolloidchem. Beih.*, 23 (1927), 406; A. FREY-WYSSLING, *Protoplasma*, 25 (1935), 261; 26 (1936), 45; 27 (1937), 372, 563; J. POLYMER *Sci.*, 1 (1946) 266.

As has been stated above, intrinsic dichroism arises from orientated molecules of a coloured substance. Fibres dyed with dyestuffs like Congo red show dichroism and according to investigations of FREY-WYSSLING and coworkers, this is due to the fact that the molecules of the dyestuff are adsorbed in an orientated position on the chains of the macromolecular component[1]. Some authors have derived information on the state of orientation of gels from measurements of dichroism[2].

c. X-ray diffraction pattern and orientation

We can discriminate two applications of X-ray research to compute orientation. One is concerned with the inner structure of the crystallites and, hence, yields information on the orientation of the molecules in the latter, the other aims at the determination of the spatial orientation of the crystallites themselves.

c. 1 *The orientation of the molecules in the crystallites*

The lattice structure of the crystallites occurring in macromolecular xerogels and gels may be investigated by the usual methods developed for the X-ray analysis of crystals. However, since monocrystals of the substance are never available in macromolecular systems, except in a few cases (proteins), the problem is more difficult here, and can, usually, not be solved with the same degree of accuracy and exactness as in ordinary crystals. By combining the outcome of X-ray analysis with stereochemical and other evidence, a fairly reliable picture of crystal structure can, nevertheless, be obtained in a number of cases. The best results may be arrived at when it is possible to examine objects with an approximately perfect biaxial orientation. The results of X-ray analysis as applied to macromolecular systems have already been often reviewed[3] and may be considered as generally known. We shall therefore only briefly refer to them here. In a number of cases the details of the crystal structure are only of secondary importance as to the problems of gel structure and the behaviour of gels. In other cases, however, some knowledge of the crystal structure is necessary for a general understanding of these items and we shall therefore especially focus attention on this side of the subject.

The general type of crystal structure in linear macromolecular substances is essentially that also shown by the simplest long chain molecules, the paraffins. The extended chain molecules, are plane zigzag structures, lying parallel to each other, and pack like flattened rods or ribbons.

The shape of the crystallites themselves is usually anisometric, their greatest dimension coinciding with the direction of the chains. They have the form of lamellae rather than that of rodlets (cf. Fig. 90 in the next section) owing to the different magnitude of the intermolecular forces in the various directions perpendicular to the chain. The lamellar planes correspond to the crystallographic planes with the densest population of those atoms or groups, which contribute most to intermolecular cohesion (as, for instance, the hydroxylgroups in cellulose).

---

[1] A. FREY-WYSSLING, *Vierteljahresschr. Naturf. Ges. Zürich*, 88 (1943) 161; O. Wälchli, *Thesis*, Zürich 1945.

[2] J. M. PRESTON, *J. Soc. Dyers and Colourists*, 47 (1931) 312; A. FREY-WYSSLING, *Jahrb. f. wiss. Botanik*, 90 (1942) 705; *Chromosoma*, 2 (1943) 473; *Die Stoffausscheidung der höheren Pflanzen*, Berlin 1935, p. 41.

[3] Cf. *e. g.* J. R. KATZ, *Röntgenspektrographie als Untersuchungsmethods*. Berlin 1934; K. H. MEYER, *The natural and synthetic High-Polymers*, New York 1942.
W. T. ASTBURY, *The principles of fibre structure*.

α. Highmolecular paraffins. In the crystals of low molecular normal paraffins and their derivatives the molecules have been shown to lie as extended chains whose terminal groups formed planes, in some cases perpendicular and in others oblique to the chain axes[1]. High molecular paraffins (polyethylene) were investigated by BUNN[2]. The long chain molecules pack in exactly the same way. The terminal groups, however, are no longer assembled in definite planes but lie at arbitrary places. It was even shown that the length of the crystallites is smaller than the average length of the chains and the latter may therefore pass from one crystalline region to another. The principal spacing in the chain axis is 2.5 Å, corresponding to the distance of the repeating unit of the zigzag chains ($CH_2 . CH_2$).

The X-ray diffraction pattern of low molecular paraffin crystals shows interesting changes when the temperature is raised, indicating that, before the melting point is reached, the lateral spacings of the chains become less differentiated and it is assumed that the chains then become free to rotate around their long axes. In macromolecular hydrocarbons the crystallites disappear beyond a certain temperature and the chains then coil up to form randomly kinked configurations. This is accompanied by the contraction of the macroscopic object if it was initially anisotropic.

The typical changes occurring in the structure of the crystals of low molecular *n*-paraffins just below the melting point are also found in the polyamides at elevated temperature[3].

β. Condensation polymers of the polyester and polyamide type[4]. The molecules of these classes of substances may be pictured as paraffin chains in which other
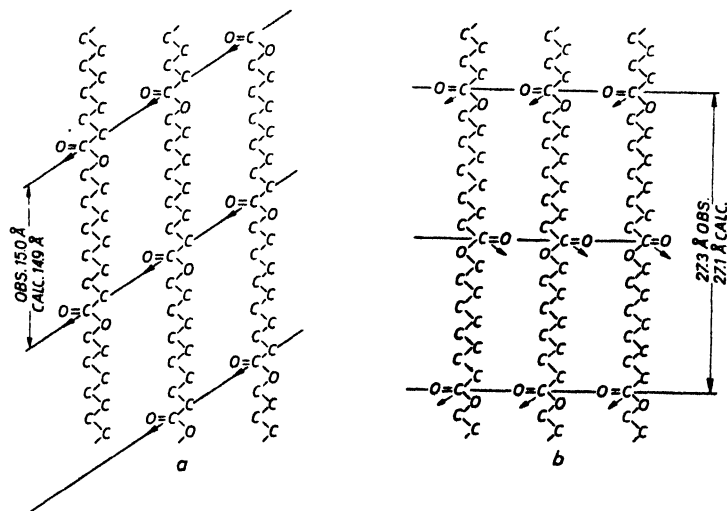
Fig. 67. Diagrammatic picture of dipole layers in which association of the polar groups takes place, a) in polyundecanoate, b) in polydecanoate (borrowed from FULLER and BAKER).

[1] A. MÜLLER, *J. Chem. Soc.*, 123 (1923), 2043; *Proc. Roy. Soc. London*, A 120 (1928), 437; MALKIN, *J. Chem. Soc.*, 1931, 2796.

[2] C. W. BUNN, *Trans. Faraday Soc.*, 35 (1939), 482.

[3] R. BRILL, *J. prakt. Chem.*, 161 (1942), 43.

[4] For a concise review, see C. S. FULLER and W. O. BAKER, *J. Chem. Ed.*, 20 (1943), 3.

groups are inserted at regular intervals. These groups are of a polar nature. The polar groups always associate to form dipole layers oblique or perpendicular to the chain axis according as the number of chain atoms of the repeating unit is even or odd. The spacings of these planes often exactly correspond to those calculated from atomic distances, sometimes they are slightly smaller than the calculated ones, indicating that the paraffin-like sections of the chains between the polar planes are probably somewhat twisted or tilted (cf. furtner on). This has been observed in certain polyamides, and occurs more markedly at high temperatures[1].

If the number of chain atoms in the repeating unit is odd, the fibre axis spacing is twice the length of the unit, since then the positions of the successive polar groups in one chain alternate from one side of the chain to the other (see Fig. 67b). The packing of the chains is less tight and the forces of cohesion between the polar groups in the polar planes is weaker than when the number of chain atoms is even (Fig. 67a). The same principle applies to polyamides. In the latter the polar forces are stronger than in the polyesters, probably owing to the possibility of hydrogen bond formation[2]. In Fig. 68 a diagram of the crystal structure of the polyamide, resulting from the condensation of adipic acid and hexamethylene diamine (Nylon polymer), is shown. The polar bonds in the polyamides are of the same type as those occurring in the linear proteins (see below).

Fig. 68. Polar layers in the crystallites of polyhexamethylene adipamide.

Polyesters and polyamides (and similar polymers, like polymeric anhydrides, polyurethanes etc.) are less waxy and harder than paraffins owing to the presence of polar layers with their strong cohesive forces. The physical properties of these substances largely depend upon the number of polar groups per unit volume and upon the magnitude of the cohesive forces in the polar layers. The melting point is the higher and the hardness the greater, according as the number of CH_2-groups in the repeating unit is smaller. Polyesters have a lower melting point and are less hard than the corresponding polyamides, owing to the stronger polar bonds in the latter; further-

[1] R. BRILL, loc. cit.
[2] A comprehensive discussion of these forces and their nature was given by W. BROSER, K. GOLDSTEIN and H. E. KRÜGER, Kolloid-Z., 105 (1943), 131; 106 (1944), 187.

more, the melting points of polymers with an odd number of chain atoms per repeating unit is lower than that of those with an even number.

If the number of polar groups becomes very large relative to the number of CH_2 groups, as for instance in the polyester from ethylene glycol and succinic acid, the interaction of the polar groups in adjacent chains (and perhaps also in the same chain) becomes so considerable as to cause the structure to seek an other more favour-
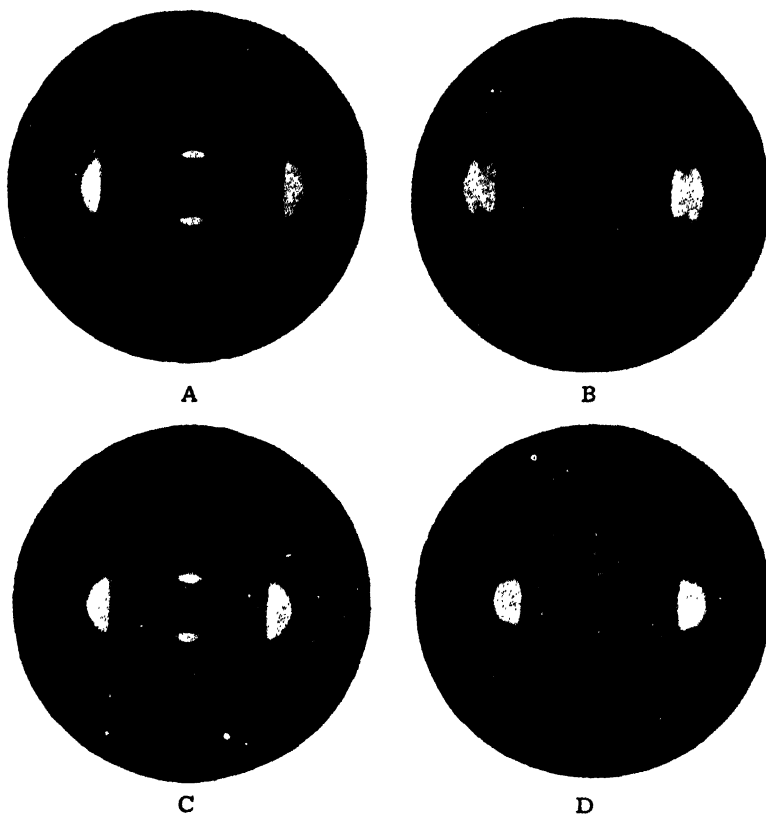
Fig. 69.   Effect of heat treatment on [X-ray diffraction pattern of fibres from a typical polyamide (polydecamethylene sebacamide). A before, B after annealing.

able type of packing. The pattern is then completely changed, the paraffin chains remaining no longer in the extended configuration. The X-ray pictures and the physical properties show abrupt changes[1].

Polyesters and polyamides are capable of giving fibres on cold-drawing of their cooled melts and some of them are used in the production of synthetic fibres, exhibiting

[1] Literature on polyesters and polyamides and their crystal structure: H. MARK, The collected papers of W. H. CAROTHERS, New York 1940; C. S. FULLER and W. O. BAKER, *J. Chem. Ed.*, 20 (1943), 3. *Polyamides*: C. S. FULLER et al, *J. Am. Chem. Soc.*, 59 (1937), 344; 61 (1940), 2575; 64 (1942), 154, 2399; *Chem. Rev.*, 26 (1940), 143; R. BRILL, *J. prakt. Chem.*, 161 (1942), 49; *Z. physik. Chem.*, B 53 (1943), 61; *Polyesters*: C. S. FULLER, *Ind. Eng. Chem.*, 30 (1938), 472.

excellent properties from a textile point of view. It is especially in this process of cold-drawing that well orientated crystalline structures particularly suitable for X-ray examination are formed. Fibres produced by cold-drawing from rapidly cooled melts (so called ,,quenched" products) often show a less well developed crystalline pattern than the same fibres after having been subjected to an annealing process (heating at elevated temperature below the melting point). An example borrowed from BAKER and FULLER is shown in Fig. 69. The fibre produced from the quenched polyamide exhibits but one equatorial interference, corresponding to only one average lateral spacing between the chains. Upon annealing this reflexion splits up into two, which correspond to two different lateral spacings in two directions. The rapid setting in of hydrogen bonding upon cooling of the melt apparently freezes portions of the chains in non-equilibrium states, which persist at the drawing process. Heating of the fibre yields a transformation into the equilibrium structure by releasing the molecules from their "frozen" positions. There is evidence, that in polyamides even careful annealing does not serve to bring about as complete order as may be reached in polyesters by a similar treatment. Owing to the much stronger polar bonds in the former, there seems to remain always some mismatching of polyamide groups. We shall see later that the ability to show "quenched" types of fibre patterns is a very general phenomen in other linear macromolecules too. It is the more pronounced, the larger are the lateral cohesive forces.

Very interesting and instructive results have been obtained by FULLER, BAKER and PAPE in systematic investigations on the X-ray behaviour and the physical properties of copolymers (i. e. polymers formed by polycondensation of mixtures of various repeating units) and methylated polyamides. Amongst other results, transformations of the X-ray pattern have been obtained upon stretching, which are similar to those observed in the α keratine- β keratine transformation in animal hairs (see below)[1]. For this subject we must, however, refer to the literature cited.

Finally, mention must be made of recent investigations of HESS and KIESSIG revealing some essentially new features as to the X-ray behaviour of polyesters and polyamides[2]. They used a special camera[3] designed for the investigation of large lattice spacings and, hence, to detect and measure interferences lying near to the point of intersection of the primary pencil of X-rays with the film. Investigating poly hexamethylene adipamide it was found that, besides the fibre periodicity of 17.4 Å already known (see Fig. 68), much longer spacings exist. They occurred in the X-ray pattern of the cold drawn fibres as well as in those of the cooled melts and were very sharp and intense (the intensity corresponding to that of the base reflexions of the ordinary paraffins).

Long fibre spacings in artificial objects were, thus far, only observed in relatively low molecular compounds, like low molecular polyoxymethylenes of uniform chain length and paraffins like $C_{30}H_{62}$ and $C_{60}H_{122}$[4]. The spacing here corresponded to the length of the individual molecules, whose terminal groups, obviously, lie in

[1] W. O. BAKER and C. S. FULLER, J. Am. Chem. Soc., 65 (1943), 1120.
[2] K. HESS and H. KIESSIG, Z. physik. Chem., 193 (1944), 196. HESS and KIESSIG could not confirm the observation of FULLER, BAKER and PAPE that the interferences of quenched melts are less sharp and less differentiated than those of the annealed products.
[3] H. KIESSIG, Kolloid-Z., 98 (1942), 213.
[4] H. STAUDINGER et al, Z. physik. Chem., 126 (1927), 425; R. KOOLHAAS and K. SOREMBA, Z. Kristallogr., 100 (1939), 47.

definite planes. The spacings found by HESS and KIESSIG, however, showed no relationship either with the length of the repeating units or with the total length of the molecules as derived from molecular weight determinations. Fig. 70 shows one of the low angle deflection X-ray photographs of polyhexamethylene adipamide in various conditions. The low angle interference shown corresponds to a lattice spacing varying between 74 Å and 88 Å. According to HESS and KIESSIG these reflections correspond to the length of the crystalline regions themselves, and this implies that the latter must have almost uniform length and that they are arranged at equal distances in the fibres. It was found, that the spacing increased upon annealing and, comparing Fig. 70-1 with Fig. 70-2, it is seen that the interference spots become sharper. Simultaneously, their intensity is enhanced. (The exposure time was 20 h in Fig. 70-1 and only 8 h in Fig. 70-2).

The crystallites from the melt showed the largest period $d = 88$ Å and it is seen from

Fig. 70.   Low angle deflection X-ray photographs of polyhexamethylene adipamide, borrowed from HESS and KIESSIG.

1. Cold drawn, stretched fibres; d = 74 Å.
2. The same after annealing at 200°; d = 79 Å.
3. Solidified melt. (not orientated); d = 88 Å.
4. Fibre drawn from melt (unstretched); d = 88 Å.

fig. 70-3 (exposure time 31 h) that they exhibit random orientation here.

Analogous results were obtained with other polyamides and polyesters, as may be seen from the following survey of d-values (Table 11).

TABLE 11

LONG FIBRE SPACINGS OF POLYAMIDES AND POLYESTERS.

	stretched fibre	melt
poly hexamethylene adipamide	74 A	88 A
poly ε aminocapronic acid	74 ,,	115 ,,
poly ω oxydecanoic acid	161 ,,	199 ,,
poly ω oxyundecanoic acid	137 ,,	185 ,,

Annealing the fibres caused a lengthening of the crystallites in any case. In polyundecanoic acid a d-value of 210 Å was reached, even surpassing that of the melt. In hexamethylene adipamide, boiling in water had a similar effect as annealing.

To demonstrate that the spacings found have nothing to do with the total chain length of the polymer, a series of hexamethylene adipamides of various molecular weight (computed from viscosity measurements) was investigated. The result is shown in Table 12.

TABLE 12

LONG FIBRE SPACINGS IN POLYHEXAMETHYLENE ADIPAMIDES OF VARYING AVERAGE CHAIN LENGTH
(HESS AND KIESSIG)

intrinsic viscosity $[\eta]_g$	0.17	0.25	0.86	1.97	1.24
molecular weight	2300	3380	11650	13130	16800
total chain length in A	179	275	905	1018	1300
long spacing in A	88	81	88	96	91

The conclusions arrived at by HESS and KIESSIG from their observations are intimately connected with theoretical points of view bound up with the deformation of these objects. Although this subject falls outside the scope of this section and belongs to § 9, it will be briefly discussed here, since it can hardly be separated from the foregoing.

From electron diffraction experiments[1] it could be derived that the dimension of the crystallites in polyhexamethylene adipamide is roughly in the order of 100 Å, which is not in contradiction to the hypothesis put forward that the long periods correspond to crystallite dimensions. Since the length of the molecular chains (Table XII) by far exceeds 100 Å in the high polymer species (M = ∞ 1000) used in practice, it must be assumed that one chain passes through several (∞ 10) successive crystallites, the latter being separated by "amorphous" regions consisting of chains in a less orderly arrangement.

Fig. 71. Diagrammatic picture of amorphous transitional region between two crystallites (HESS and KIESSIG).

By virtue of some yet unknown natural cause (HESS and KIESSIG speak of the vibrational energy of the chains) long chain-molecules are not capable of taking part in the formation of lattice order over their entire length. In the process of crystallization amorphous sections remain between the crystallites (Fig. 71), but these crystallites tend to assume a uniform length, which, however, varies with the temperature of crystallisation.

For a given temperature there would be an equilibrium size of the crystallites.[2] A diagram showing the structure of an orientated fibre and the periodicity P due to the crystallite length is shown in Fig. 72.

In connexion with these ideas, some details of the X-ray diffraction photographs are of interest. The photographs of the orientated stretched fibres (Fig. 70-1) show a considerable lateral broadening of the interference spots which is considerably

[1] M. v. ARDENNE, E. SCHIEBOLD and F. GÜNTHER, Z. Physik, 119 (1942), 363.
[2] C.f. also E. M. FRITH and R. F. TUCKETT, Trans. Faraday Soc., 40 (1944) 251.

reduced upon annealing. In contrast to this lateral extension of the low angle diffraction spots, the ordinary interferences of the well orientated fibres (not shown in Fig. 70) are sharp and narrow in conformity with the nearly perfect axial orientation of the crystallites. In this respect HESS and KIESSIG refer to a discussion of SCHIEBOLD[1], showing that a linear lattice consisting of threadlike parallel particles, each individually exhibiting the same lengthwise periodicity, but so arranged that no lateral periodicity comes into being, will exactly yield such a lateral broadening of the meridional reflexions. The lack of lateral periodicity might for instance be due to unequal lateral dimensions of the crystallites. Upon annealing the crystalline regions are completely broken down, and simultaneously rebuilt to equilibrium length, now also assuming more uniform lateral dimensions, resulting in a better lateral periodicity.

According to HESS and KIESSIG, the process of cold drawing is also accompanied by a breakdown of existing crystallites and the formation of new shorter ones of unequal cross section, the long range periodicity dropping from 88 in the cooled melt to 74 Å in the orientated fibres.

It would seem that the flexibility and elasticity of the fibres of these synthetic polymers, combined with a high tensile strength, can be more or less understood from the picture developed. The latter property may be associated with the long primary valence chains extending through a series of crystallites (where they are tightly bound against each other and prevented from slipping) and further with the approximately equal length of the chains in the amorphous

Fig. 72. Diagrammatic picture of an orientated fibre and the long range period P. (HESS and KIESSIG).

regions[2]. The flexibility and elasticity may be associated with the amorphous regions. The irregularly distributed polar groups there lie, so to speak, imbedded in a matrix of paraffin and considerable internal movements and displacements will be possible without a destruction of the structure.

It is of considerable interest, that polyhexamethylene adipamide, prepared by condensation in a solvent like xylenol and then precipitated in powder form, does not show long range periodicity, neither do fibres spun from a xylenol solution in an alcohol bath (HESS and KIESSIG). The fibres obtained have a low tensile strength and are brittle. These objects are obviously formed in the gel condition at ordinary temperature and then exhibit less favourable properties. These facts are important with regard to regenerated cellulose.

Long range spacings have not been observed in either natural or artificial cellulose fibres, nor in extended rubber.

γ. *Polymers formed by polymerisation*[3]. So far we have been concerned with what may be termed condensation polymers. In this section, we shall focus attention

---

[1] E. SCHIEBOLD, *Kolloid-Z.*, 69 (1934) 281.

[2] Upon loading the fibre, these chains will simultaneously reach the state of maximal extension. If the amorphous chain sections are of unequal length, they will break one by one instead of simultaneously and thus the tensile strength at break will be lower. Cf. P. H. HERMANS, *Kolloid-Z.*, 108 (1944) 177.

[3] The writer has followed the survey given by C. S. FULLER and W. O. BAKER, *J. Chem. Ed.*, 20 (1943) 3.

on the products formed in the polymerisation of vinyl derivatives (cf. Chapter II). This process may be generally represented by the formula:

$$n. \quad \overset{R_1}{\underset{R_2}{>}} C = CH_2 \rightarrow - \overset{R_1}{\underset{R_2}{C}} - CH_2 - \overset{R_1}{\underset{R_2}{C}} - CH_2 - \overset{R_1}{\underset{R_2}{C}} - CH_2 -$$
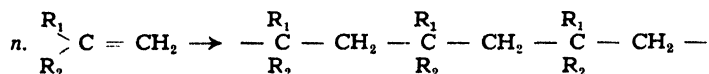
where R_1 and R_2 may be H, Cl, CH_3, OH, $COOCH_3$ etc.

Paraffin chains bearing groups laterally attached to the chain at every other C-atom are formed in this process. Consequently, if these groups are of such a nature as to interact with oneanother strongly, they will tend to give permanently quenched or amorphous glass-like masses. This is particularly true when R_1 differs from R_2,
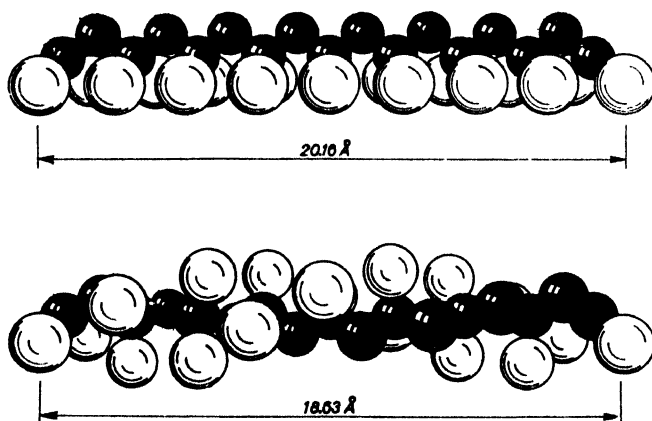
Fig. 73. Extended form of the polyisobutylene chain (upper part) compared with probable spiral form (lower part), exhibiting a fibre period of only 18.6 A. (after FULLER and BAKER).

as in polyvinylchloride, for example, where R_1 = H and R_2 = Cl and poor X-ray fibre patterns result. On the other hand, where definite bonds can form between the R-groups as in polyvinylalcohol, the tendency is to favour a definite crystalline lattice. Orientated polyvinylalcohol fibres show a very sharp X-ray pattern, indicating a high degree of order. In this case R_1 = H and R_2 = OH. Hydrogen bonding occurs between the OH-groups. The fibre period is 2.52 Å, corresponding with the distance between two successive OH groups in agreement with the ordinary zig-zag backbone structure of the carbon chain.

A different behaviour is to be expected if R_1 and R_2 are nonpolar groups with only weak cohesive forces, like for instance methylgroups as in polyisobutylene. A high degree of order is possible in the extended chains as shown by Fig. 73 (upper part).

The observed fibre period is 18.6 Å instead of 2.52 Å, indicating that repetition in the chains is postponed until many monomeric residues have been traversed. A way to achieve this is to have the chain twist (Fig. 73, lower part). Eight units then reach the length of 18.6 Å, whereas in the extended form their length would be 20.2 Å. Thus the molecules appear to be helices in which repetition of the structure

occurs at every ninth methylgroup. It is assumed that right and left helices exist together in the crystalline structure.

A similar structure is assumed to exist in the crystallites of rubber and other rubberlike hydrocarbons[1].

An outstanding feature of rubber at ordinary temperature and of other polymers exhibiting rubberlike properties in a certain range of temperatures (cf. Chapter XIII) is that they are amorphous, and crystallize upon cooling, and also upon stretching to a sufficient degree of extension. Crystallites are then formed at the cost of the amorphous substance. The crystallisation of rubber, which was discovered by KATZ in 1928[2], has recently been quantitatively investigated by FIELD and GOPPEL[3].

It will be clear that by relatively simple chemical changes of the groups R_1 and R_2 one can produce materials which are either rubberlike in character, as in the case of polyisobutylene, or which are more crystalline at ordinary temperatures and represent harder, tougher plastics. The fundamental properties operating to cause these outstanding differences are, of course, the magnitude of the forces acting between the substituting atoms or groups on adjacent chain molecules.

δ. *Cellulose.* Amongst the linear polymers, cellulose is perhaps the case in which the most detailed information on crystal structure has been obtained. Four different
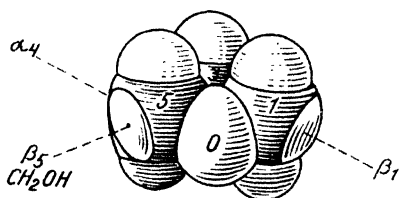
Fig. 74. STUART-model of the β glucose ring system in cellulose with only the hydrogen atoms attached to it. The C-H bonds are perpendicular to the ring plane.
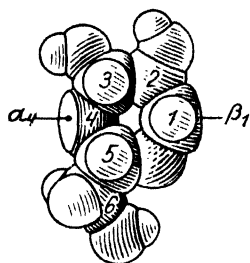
Fig. 75. The model of Fig. 74 viewed from the top and with 2 OH groups and the CH_2OH group attached to it. 1, 3 and 5 are hydrogen atoms.

lattice modifications have thus far been discovered. Two of these, cellulose I and cellulose II are those occurring in ordinary native and regenerated fibres respectively and are the most important ones.

The most recent investigations on the crystal structure of cellulose are those by MEYER and MISCH, GROSS and CLARK, KIESSIG, KUBO, SCHIEBOLD and PEIRCE[4].

[1] J. R. KATZ, *Kolloid-Z.*, 36 (1925), 301; 37 (1925), 19; H. MARK and G. V. SUSICH, *Kolloid-Z.*, 46 (1928), 11; W. LOTMAR and K. H. MEYER, *Monatsh.*, 69 (1936), 115; E. SAUTER, *Z. physik. Chem.*, 36 (1937), 405; H. MORSS, *J. Am. Chem. Soc.*, 60 (1938), 237; C. B. J. CLEWS, *Rubber Chem. Techn.*, 12 (1939), 19; L. MISCH and A. J. A. VAN DER WIJK, *J. Chem. Phys.*, 8 (1940), 127; R. BRILL and F. HALLE, *Naturwiss.*, 26 (1938), 12.

[2] J. R. KATZ, *Naturwiss.*, 13 (1928), 410.

[3] J. E. FIELD, *J. Appl. Phys.*, 12 (1941), 23; J. M. GOPPEL, *Appl. Sci. Res.*, A 1 (1947), 3.

[4] K. H. MEYER and L. MISCH, *Helv. chim. acta*, 20 (1937) 232; S. J. GROSS and G. L. CLARK, *Z. Kristallogr.*, 99 (1938) 357; H. KIESSIG, *Z. physik. Chem.*, B 43 (1939) 43, 79; T. KUBO, *ibid.*, A 187 (1940) 297; E. SCHIEBOLD, *Kolloid-Z.*, 108 (1944) 248; F. T. PEIRCE, *Trans. Faraday Soc.* (1946).

In all modifications the same type of fibrillar structure recurs, and the extended chains lie parallel to each other. Only the lateral arrangement of the chains (or perhaps also the configuration of the ring systems in the glucose units) is slightly different in the various modifications.

The repeating units in the chain direction consist of two glucose residues inter-linked by a β-glucosidic bond. A stereochemical discussion of the configuration of the molecule in connexion with its packing in the crystallites was recently given by the present author et al[1]. It was concluded that the 6-membered pyranose ring system of the glucose residues assumes the so called armchair form of the cyclohexane ring. In this configuration the C-H-bonds of the five hydrogen atoms directly bound
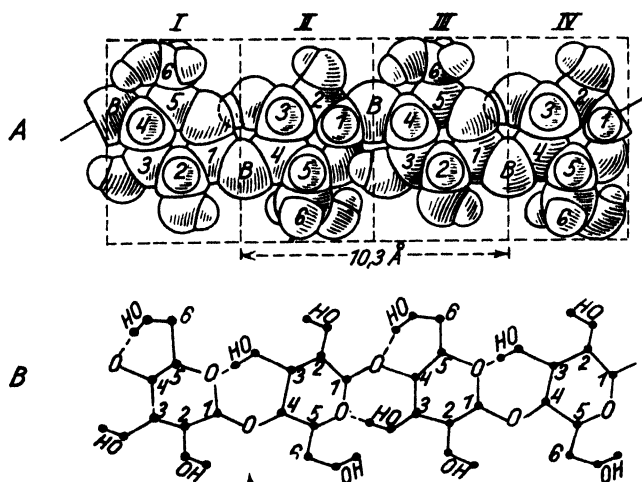
Fig. 76

a) STUART model of the cellulose chain. Notice the zigzag position of the bridge oxygen atoms B. All the C - H valencies are directed perpendicular to the plane of drawing; the position of all the hydroxylgroups is sideward.
b) Diagram of the corresponding valence frame. Hydrogenbonds indicated by dots.

to the ring carbon atoms are all perpendicular to the ring plane as shown by the STUART model in Fig. 74. In Fig. 75 the same model, with the OH-groups and the CH_2OH group attached to it, is shown viewed from the top.

Finally Fig. 76a shows four glucose residues forming a section of a cellulose chain, having dimensions consistent with the X-ray data. The lines connecting the 0 atoms of the glucosidic bridges (B) form a zig zag, and the whole molecule has the symmetry of a diagonal screw. Its shape is rather that of a flat ribbon than cylindrical The top and the bottom planes of the ribbon are occupied by hydrogen atoms, and the hydroxylgroups are placed at the smaller sides of the ribbon. This particular. molecular configuration can only be reached if the glucose residues are in the β-glucosidic armchair configuration. In Fig. 76b it will be seen that intermolecular hydrogenbonds may occur between the OH groups at the carbon atoms 3 and

[1] P. H. HERMANS, J. DE BOOYS and CHR. J. MAAN, Kolloid-Z., 102 (1943) 169.

the ring oxygen atom, which stabilize the straight configuration of the chain.

The ribbon shaped molecules are packed in the crystallites as diagrammatically indicated in Fig. 77. The hatched areas are the regions where the hydroxyl groups lie. Another picture, showing the approximate position of the main valence frame in the elementary cell is reproduced in Fig. 78.
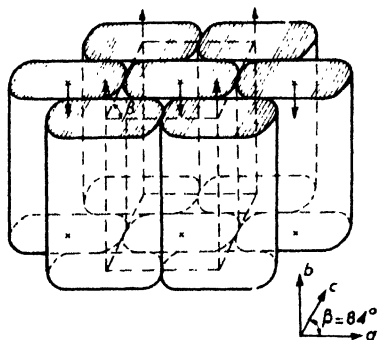
Fig. 77. Diagrammatical picture of the packing of the cellulose chains in the crystallites. The prism marked with broken lines is the elementary cell. The regions occupied by the hydroxyl groups are hatched. (The arrows indicate the alternating polarity of the chains.
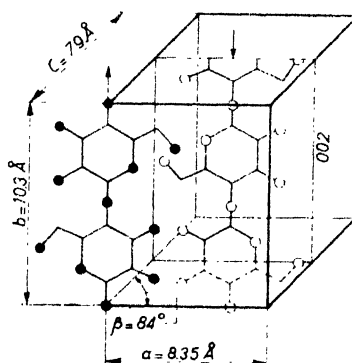
Fig. 78. Approximate position of the main valence frame of the cellulose chains in the elementary cell.

Numerous hydroxyl groups (six per repeating unit) being present in the cellulose chain, the lateral cohesion in the crystallites is very strong, and extends over the whole length of the crystalline regions, which, hence, have to be considered as very rigid and stable structures. In these lateral bonds hydrogen bonding plays a considerable part. The position of these hydrogen bonds has been recently discussed at length by SCHIEBOLD (loc. cit.)[1]. Owing to the strength of these lateral bonds the melting point of cellulose is very high and cannot be reached without previous decomposition. The insolubility of cellulose in water has no doubt the same cause.

Since the details of the crystal structure are of very little importance with regard to the structure and the behaviour of cellulose gels, we shall forbear from discussing them. Neither the lattice structure nor the amount of crystalline substance in cellulose has, as yet, been found to be affected by ordinary mechanical deformation. The orientation of the crystallites changes, but it would seem as though the crystallites themselves remain, as a rule, unaffected. In all regenerated cellulose fibres the percentage of crystalline substance is astonishingly constant, a fact which remains to be explained[2].

The suitability of cellulose to form fibres may be pictured as being due to the combined effects of the large cohesive forces in the crystallites on the one hand, and the occurrence of amorphous

---

[1] See also the recently appeared papers by F. T. PEIRCE, Trans. Faraday Soc., 42 (1946) 545, 560 where also a general improvement of the picture of the crystall structure of cellulose accounting for some thus far unexplained features in the X-ray pattern is given.

[2] P. H. HERMANS and A. WEIDINGER, J. Appl. Phys., 19 (1948) 491.

fringelike portions interconnecting the latter on the other hand. Flexibility and elastic properties will again have to be associated with the amorphous portion of the fibre. These properties are, however, only observable if a sufficient quantity of water is present, in order to reduce the inter-molecular cohesive forces in the amorphous regions.

If regenerated cellulose fibres are formed by the deformation of swollen gels, there is no reason to expect that the molecular fringes between the crystalline regions should have equal length. This may be the reason for the relatively low tensile strenght of these objects (particularly in the wet state) as compared to that of native fibres like ramie and cotton. In the latter, formed by the action of life, a more regular structure may be expected.

A phenomenon analogous to the effects of annealing in the polyamides, shown in Fig. 69, is observed when cellulose gels freshly prepared from viscose (cellulose xanthate solutions) are boiled in water. The two main equatorial reflexions A_3 and A_4 of regenerated cellulose are not separated in the freshly prepared gels, but become so on boiling. A certain recrystallisation seems to occur also when regenerated cellulose swollen in glycerol is heated at 250°[1].

Ceilulose esters and ethers also give fibrillar crystal structures of the same type. The cohesive forces between the chains are, however, weaker than in cellulose. Cellulose triacetate and triethyl-cellulose, for instance, show a melting point, but melting can not be accomplished without a marked breakdown of the chains. According as the substituents themselves are larger and their polarity decreases, the melting point of the derivatives becomes lower, and their plasticity increases. This becomes clearly apparent if the properties of the triesters of

Fig. 79. Diagrammatic representation of ordered (upper) and disordered (lower) systems of parallel cellulose chain systems, with the type of X-ray pattern each produces.

cellulose of the homologous series of the normal fatty acids are compared[2].

Cellulose esters have been shown to exhibit the phenomenon of increasing local order upon annealing to a very marked degree[3]. The typical change of the X-ray pattern is diagrammatically shown in Fig. 79, borrowed from FULLER and BAKER In the two cases represented in this picture systems of parallized chains are considered. In the upper part of the figure the chains are ordered in three dimensions and arranged to from a true crystalline lattice. The orientation of the crystallites themselves is supposed to be random. The X-ray pattern consist of a number of sharp DEBYE-SCHERRER rings. In the lower part of the figure the lateral order of the chains

[1] P. H. HERMANS, *Contribution to the Physics of Cellulose fibres*, Amsterdam-New York 1946.
[2] HAGEDOORN and MOELLER, *Cellulosechemie*, 13 (1931) 29.
[3] C. S. FULLER and W. O. BAKER, *J. Am. Chem. Soc.*, 64 (1942), 776; *J. Chem. Ed.*, 20 (1943), 3.

is supposed to be disturbed in that they are more or less lengthwise spiraled. The X-ray pattern then consists of broader less sharp rings. Annealing causes an ordering of the chains in each bundle by allowing them to rotate into their lattice equilibrium positions.

Phenomena of this kind play an important part in cellulose plastics. Recently they have been extensively studied in connection with physical properties of fibres, films and molded products by Baker[1]. This author also holds the view that a number of cases, where the existence of various lattice modifications of cellulose derivatives have been supposed to exist on account of different X-ray diffraction patterns, should rather be regarded as representing states of different degree of order.

Hess et al[2] have shown that crystalline order in macromolecular systems may be disturbed by strong mechanical disintegration. They showed that cellulose fibres upon milling in a vibrating ball mill of special design entirely lost their crystalline character in one hour, the X-ray photograph only showing an amorphous ring. Upon wetting with cold water recrystallisation occurred at once and the X-ray pattern appearing was that of cellulose II (hydrate cellulose). This example shows that in polar substances like cellulose a swelling agent may give rise to a similar effect as heating does in other cases.

r. Proteins. This very interesting class of macromolecular substances occurring in nature is characterized by the fact that the lattice structure itself may show considerable changes upon mechanical deformation of the macroscopic objects and, according to a hypothesis put forward by Astbury, this would be intimately connected with a change in the configuration of the chains.

Proteins may be classified in two groups, *fibrillar* proteins which are, as a rule, insoluble, and soluble *globular* proteins. Silk fibroin, keratin (the substance of animal hairs, horn, nails) and collagen (the substance of animal hide and tendons), belong to the former, ovalbumin, haemoglobin and a number of seed proteins belong to the latter. The muscle protein myosin, though soluble in salt solutions, is also considered as a fibrillar protein on account of its similarity with keratin[3]. The globular proteins are considered as being metastable, they easily transform into an other form, intimately connected with fibrillar proteins (denaturation)[4].

The X-ray diffraction pattern of fibrillar proteins is, as a rule, less sharply defined and poorer in reflections than that of cellulose, but have, none the less, been obtained as characteristic fibre-diagrams. (Silk fibroin[5] (Cf. Fig. 85 next section) and feather keratin are favourable exceptions[6] with regard to the differentiation of their X-ray diffraction pattern). Apart from an ordinary spacing in the direction of the fibre axis, of the order of 10 Å, a long range periodicity has been observed in many fibrillar proteins, e. g. in various specimen of collagen (tendon of the rat and kangarroo tail, beef *tendo Achilles*, cornea and ligament, human skin) keratins

[1] W. O. Baker, *Ind. Eng. Chem.*, 37 (1945), 246.

[2] K. Hess, H. Kiessig and J. Gundermann, *Z. physik. Chem.*, B 49 (1941), 64.

[3] On myosin, cf. H. H. Weber, *Pflügers Arch.*, 235 (1934) 206; K. H. Meyer and L. E. R. Picken, *Proc. Roy. Soc. London*, (B) 124 (1937) 29.

[4] e. g. W. T. Astbury *et al*, *Biochem. J.*, 29 (1938) 2351; *Nature*, 142 (1938) 33.

[5] R. Brill, *Lieb. Ann.*, 434 (1923) 204; O. Kratky and S. Kuriyama, *Z. physik. Chem.*, B 11 (1931) 363; C. Trogus and K. Hess, *Biochem. Z.*, 260 (1936) 376.

[6] W. T. Astbury and H. J. Woods, *J. Text. Inst.*, 23 (1932) T. 17.

(feather keratin, human hair, porcupine quill) myosin and others[1] showing interesting relations with the modern concepts on long range periodicity in the distribution of certain side chains in the protein molecule.

Many globular proteins, on the contrary, some of which form well shaped crystalls, yield very different diagrams[2] exhibiting all the characteristics of those of ordinary monocrystals, though, obviously, having a very complicated structure. The dimensions of the elementary cells of these lattice structures lies in the same order of magnitude as that of the above mentioned long range periodicities of the natural fibrillar proteins. The amount of substance contained in the elementary cell is of the order of magnitude

Fig. 80.   X-ray photographs of a) α-keratin and b) β-keratin.

of the molecular weight of the proteins concerned as determined in the ultracentrifuge by THE SVEDBERG[3]. They hence seem to contain one protein molecule[4]. Several observations revealed an intimate relationship between the structure of the globular and that of the fibrillar proteins[5]. It would seem that the large elementary cells of globular proteins have smaller units in common with those occurring in fibrillar proteins. In the former these units are orientated in a different, more complicated pattern and with various spatial orientation, whereas in fibrillar proteins they are all parallelized. Upon denaturation (as e. g. upon heating) the chains, folded to a certain pattern in the native globular proteins, are straightened out and may then be parallelized to a more or less extent, to form a protein, exhibiting fibrillar

---

[1] R. B. COREY and R. W. G. WYCKOFF, J. Biol. Chem., 114 (1936) 407; G. L. CLARK, E. A. PARKER, J. A. SCHAADE and W. J. WARREN, J. Am. Chem. Soc., 57 (1935) 1509; O. KRATKY, A. SEKORA, and H. H. WEBER, Naturwiss., 31 (1943) 91; O. KRATKY and A. SEKORA, J. makromol. Chemie, 1 (1943) 113; R. S. BEAR, J. Am. Chem. Soc., 66 (1944) 1296, 2043.

[2] J. D. BERNAL and D. CROWFOOT, Nature, 133 (1934) 794; D. CROWFOOT, Proc. Roy. Soc. London A, 164 (1938) 580; J. D. BERNAL et al, Nature, 141 (1938) 521.

[3] THE SVEDBERG, Proc. Roy. Soc. London, A 170 (1939) 40.

[4] The ingeneous, but speculative so called cyclol model of protein structure put forward by D. M. WRINCH (e. g. Phil. Mag., 25 (1938) 705, is not consistent with the X-ray data according to D. RILEY and J. FANKUCHEN, Nature, 143 (1939), 648. Cf. also the critisism by L. PAULING and C. NIEMANN, J. Am. Chem. Soc., 61 (1939) 1860; A. NEURATH, J. Phys. Chem., 44 (1940) 296.

[5] J. D. BERNAL, J. chim. phys., 35 (1938) 179; Nature 143 (1939) 663; W. T. ASTBURY, S. DICKINSON and K. BAILEY, Biochem. J., 29 (1938) 2351.

structure [1]. In the latter, typical fibre diagrams may also occur. Animal hairs (and also horn) exhibit the so-called α-keratin diagram (fig. 80a). Hairs stretched in steam for about 100% yield the β-keratin diagram (Fig. 80b). This has been interpreted by ASTBURY assuming that the polypeptide chains are folded according to a certain regular pattern in α-keratin and that, as a result of the stretch to which the hair is subjected, the chains are straightened out to form another regular lattice arrangement [2]. Fig. 81 shows a diagrammatic picture of the folded protein chain as occuring in α-keratin according to ASTBURY.

The side chains R protrude perpendicular to the plane of folding on both sides of the planes alternately. With STUART atomic models it can be shown that the chains on both sides are closely packed against each other, forming triangular columns [3]. The fibre period of ∞ 5 Å would agree with the lateral distances of the side chains in this arrangement and thus be determined, so to speak, by their thickness.

In β keratine, the polypeptide chains are extended and straightened but for the normal zig-zagging of the chain (Fig. 82a).

Fig. 81. Folding of the polypeptide chains in α-keratin according to ASTBURY and BELL.

The side chains are perpendicular to the plane of the paper, protruding alternately above and below that plane.

The transformation of α keratin of animal hair into β keratin by stretching is facilitated by adequate previous treatment such as heating in steam or the action of certain chemicals. These probably loosen the bonds between the ajacent side chains of the triads.

Objects consisting of β keratin yield an interference on the meridian of the diagrams, indicating a periodicity of 3,5 Å in the direction of the fibre axis. This corresponds to the length of the repeating RH . CO . NH units of the extended polypeptide chain [4]. In fig. 82b a diagram of the β keratin lattice spacings is given, viewed in a direction perpendicular to Fig. 82a. The lateral distances of the chains in the direction

[1] For further details see e. g. W. T. ASTBURY and W. A. SISSON, *Proc. Roy. Soc. London*, A 150 (1935) 535; H. NEURATH, *J. Phys. Chem.*, 44 (1940) 296; R. BRILL, *Naturwiss.*, 29 (1941) 220; J. D. BERNAL, *Nature*, 143 (1939) 663; W. T. ASTBURY, *Nature*, 142 (1938) 33.

[2] C.f. the recently somewhat modified theory given by W. T. ASTBURY et al, *Nature*, 147 (1941) 696; *J. chem. Soc. London*, 1942, 337.

[3] Lateral bonds of the hydrogen bond type or heteropolar bonds between free NH_2 and COOH groups of adjacent sidechains are assumed to interconnect the side chains in this condition.

[4] In the collagen group a configuration with a spacing of 2.7 — 2.9Å occurs. For a discussion of this subject and a survey on the very interesting behaviour of collagen c.f. W. T. ASTBURY, *J. Int. Soc. Leather Trade Chemists*, 24 (1940) 69; cf. A. KÜNTZEL and F. PRAKKE, *Biochem. Z.*, 266 (1933) 243.

of the side chains is about 10 Å (Fig. 82b) and that in the direction perpendicular to it (Fig. 82a) 4.5 Å. The latter was called the "backbone spacing" by ASTBURY. The interferences corresponding to these spacings appear on the equator of the X-ray photographs.

Extended muscle fibres and artificially produced protein fibres have a similar structure.

We shall have to forbear from going into further details on the very important and interesting subject of the X-ray spectrography of proteins in relation to their
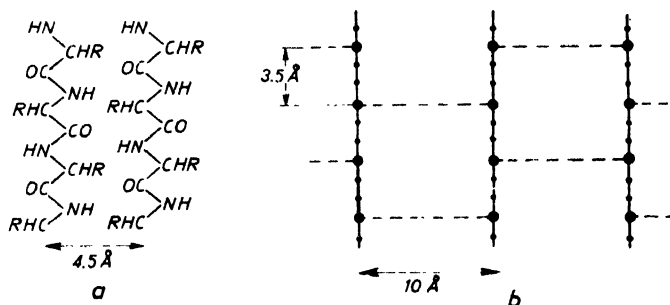
Fig. 82.   Structure of β keratin.

structure and their behaviour, this being too specialized a subject for which we must refer to the literature.

Proteins are of outstanding importance, being an intrinsic part of the material chosen by nature as the background or molecular frame of life. The kind and sequence of the side chains allow of a great many variations and seems to be determined by the function of the protein and the species of the organism. The most organized form in which a given polypeptide chain may occur seems to be the globular form. Upon denaturation (by the action of heat or chemicals) the highly organized globular configuration is destroyed and transformed into a mass of chain molecules, which may either remain an amorphous aggregate devoid of any regular structure, or assume, more or less ordered configurations like those occurring in the linear proteins. This again may occur either spontaneously, or by artificial means like stretching or spinning. Such systems behave in many respect in a similar way as other systems consisting of linear macromolecules, They are, however, characterized by the fact that allowance has to be made for a deformability of the chains within the crystallites themselves, which sometimes may take place without destruction of lattice order, in that different lattice modifications are connected with different degrees of extension of the molecular chains. On p. 606 we have seen that this is not an exclusive property of proteins. Similar phenomena have been observed by BAKER and FULLER in synthetic poly-mers.

It should be emphasized that R. S. BEAR[1] has recently put forward considerable evidence that the views of ASTBURY et al on the correlation between the macroscopic extension of protein fibres and the extension of molecular chains require a revision in many respects. This evidence is based on a very careful investigation of X-ray

[1] R. S. BEAR, J. Am. Chem. Soc., 66 (1944), 1296, 2043.

patterns taken with high resolution (small angle diagrams) which in a number of fibrillar proteins exhibit very interesting large periods in the directions of the fibre axis as well as in the lateral directions. These periods agree well with the results of observations with the electron microscope in various instances, as particularly in collagen fibres. The large longitudinal periods found (640 Å in collagen, 95—198 Å in various kinds of keratin) show correlations with the mechanical and chemical previous treatment of the fibre which seem to be in contradiction to ASTBURY's interpretation. The well known equatorial sickles seen on the ordinary "large angle" X-ray diagrams of the keratins were shown by BEAR to have a fine structure and to consist of a great number of separate interferences, equatorial as well as layer line reflections, due to some complicated structural pattern of the protein, and lying close together. Their interpretation is, hence, far less simple than thus far believed.

According to BEAR it is more likely that in the fibrillar proteins regions of higher and less pronounced order ("crystalline" and "amorphous" regions) alternate in the logitudinal direction in some regular manner. Many details of the X-ray patterns might be due to the former only, the extensibility and retractive power might well be correlated with the latter. According to these ideas the molecular mechanism of the deformation of fibrillar proteins would be more similar to that now assumed to be associated with the simpler linear macromolecular systems in general.

c. 2. Orientation of the crystallites

If a gel contains particles of a crystalline nature, its X-ray diffraction pattern will reveal whether the spatial orientation of the crystallites is at random or whether preferred orientation occurs. (cf. Volume I, Chapter I). The X-ray method has contributed a great deal to investigations on orientation. As a matter of fact it is the

a   b
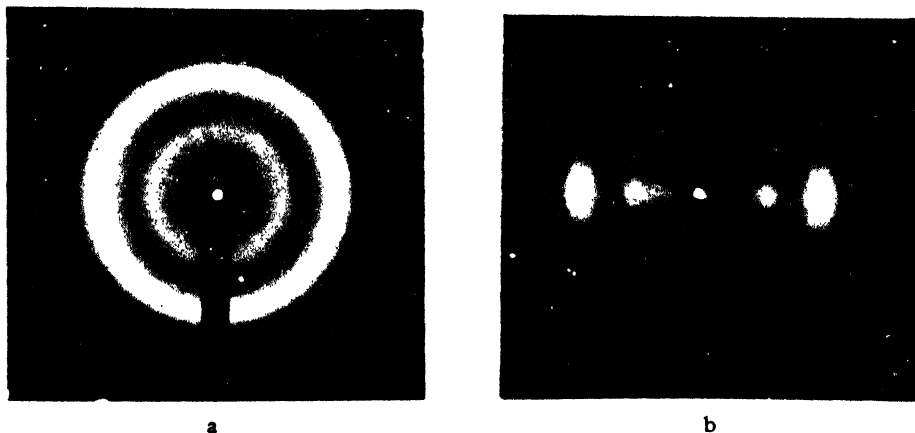
Fig. 83.   X-ray photographs of a) an isotropic, b) a highly oriented artificial cellulose filament.

only method, thus far available, from which absolutely reliable information on the state of orientation can be obtained. Its results are, however, exclusively confined to the orientation of the crystalline portion of the gels.

It is not the proper place here to enter into the details of either the general

theory or the methods of X-ray examination, and we may in this respect refer to the literature cited in Vol.¹I, Chapter I. Uniaxial orientation may, as a rule, be computed

a   b
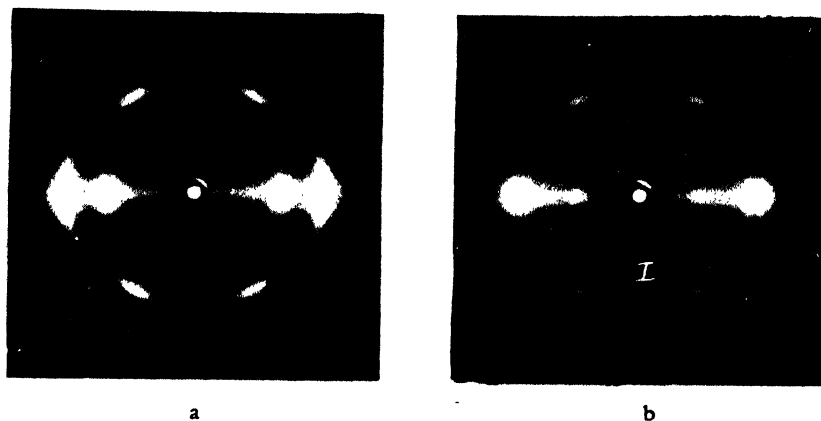
Fig. 84. X-ray photographs of a) native ramiefibres, b) mercerised ramie fibres (both highly oriented).

from X-ray photographs taken with the beam of X-rays perpendicular to the axis of orientation. In the case of biaxial orientation photographs will have to be taken

Fig. 85. X-ray photograph of natural silk.

in different directions. A very good survey of the different possible cases has been given by SISSON¹ for cellulose in natural and artificial objects.

Orientated crystallites have been revealed in a great many natural and artificial objects. Fig. 83a shows X-ray photographs of an isotropic cellulose filament and b a photograph of the same object after an extension of about 100 %. Both pictures were obtained with the beam of X-rays perpendicular to the fibre axis. It is seen, that the „DEBYE-SCHERRER diagram" appearing in Fig. 83a and indicating randomness of orientation changes into a "fibre diagram", the circular interferences being now contracted to relatively narrow sickles.

Fig. 84a shows the photograph of a china grass (ramie) fibre, taken in the native condition, showing a very high degree of orientation. Fig. 84b shows a picture of the same fibre after mercerisation (care was taken that the disorientation accompanying the treatment with the mercerizing liquor, was counteracted by

¹ W. A. SISSON, *J. phys. Chem.*, 40 (1936), 343; 44 (1940), 513; *Contr. Boyce Thompson Inst.*, 7 (1938), 230, 381; *Chem. Rev.*, 26 (1940), 187.

a subsequent stretching of the fibre). The difference in the location of the inter-ferences as compared to Fig. 84a is due to the fact that the crystalline modification of native fibres (cellulose I) is somewhat different from that in mercerized and regenerated ones (cellulose II). Fig. 85 shows the X-ray diffraction pattern of silk fibroin (natural silk filaments) revealing a high degree of orien-tation[1]. In fig. 86 the diagram of muscle fibres (at rest) is reproduced, showing some crystalline orientation next to an "amorphous ring" due to non -or very imperfectly-orientated components. Fig. 87 shows the photograph of muscle fibres of *Mytilus edulis* dried after stretching according to PICKEN, revealing a very good orientation. In Fig. 88 finally, X-ray diagrams of rubber in the isotropic and in the stretched condition are re-produced. In the former case no trace of a crys-talline portion is observed. The picture is essentially that of a liquid, only showing an amorphous ring.

Fig. 86. X-ray photograph of muscle fibres at rest after MEYER and PICKEN.

Upon stretching orientated crystallites are formed at the cost of the amorphous components and a crystalline diffraction pattern appears[2]. Other rubberlike substances like gutta-percha and balata often show similar phenomena.
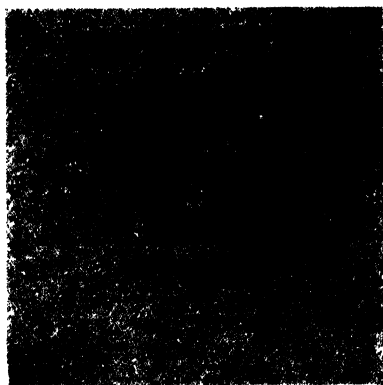
Fig. 87.   Muscle fibres of mytilus, dried in a stretched condition.

It is interesting to note that, in rubberlike substances, the crystallites formed upon stretch-ing are almost perfectly orientated in the direction of stretch from the very moment that they appear, in contrast to cellulose (and other ex-amples) where the crystallites, already existing before the deformation sets in, are gradually orientated. The difference is, however, less fundamental than it appears at first sight. G. L. CLARK has found that the X-ray spectrum of moderately vulcanized rubber, moderately extend-ed at room temperature and subsequently frozen, shows the presence of crystallites with a less perfect orientation too, which have been formed as a result of the freezing. These crystallites have been formed by the less well orientated fringes which occurred in the extended sample, and which only crystallize at a lower temperature. The spectrum of a rubber sample of this kind is quite comparable with that of a cellulose fibre as regards the type of orientation.

These few examples may serve to illustrate the widespread applicability of X-ray examination to structural problems concerned with orientation.

K. H. MEYER and L. R. E. PICKEN, *Proc. Roy. Soc. London* (B) 124 (1937) 29.
    [2] This phenomenon was first discovered by J. R. KATZ, *Naturwiss.*, 13 (1925) 410; cf. E. A. HAUSER and H. MARK, *Kolloidchem. Beih.*, 22 (1926); J. M. GOPPEL, *Appl. Sci. Res.*, 1 (1947) 3.

Most X-ray investigations on orientation have, as yet, been rather of a qualitative than of a quantitative nature. The reason of this may be, that the methods concerned with quantitative work are rather laborious and often complicated. Their principles
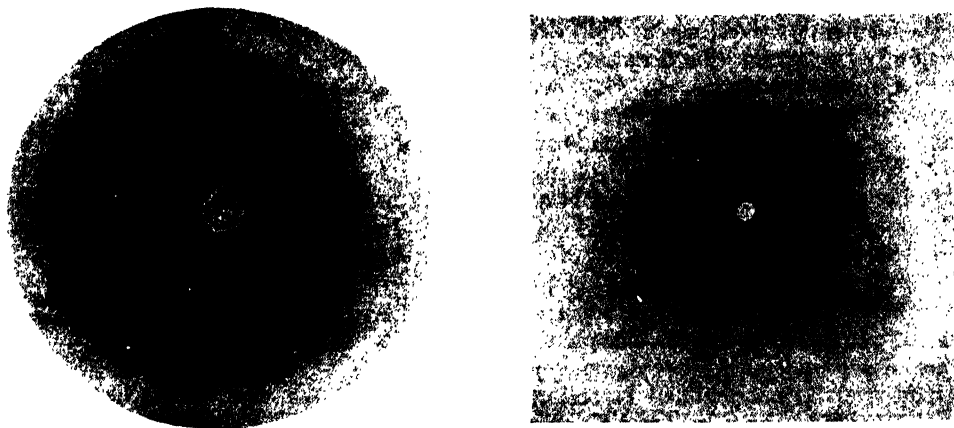
Fig. 88.
X-ray photographs of rubber a) isotropic, not stretched, b) stretched several hundred per cent.

have already been given by POLANYI[1], but have not, however, until recently been actually applied to special cases, particularly cellulose. Since their application seems to be of paramount importance for a further development of the theories of deformation, and also in other respects, we shall here briefly outline some of the outstanding points with reference to cellulose fibres, these being the only objects thus far investigated in some detail. The principles concerned simultaneously apply, however, also to other cases of uniaxially oriented objects. Fig. 89 shows a diagrammatic cross-section through the elementary cell of the lattice of cellulose II as derived from X-ray analysis (§ 8c. 1δ). The fibre axis is perpendicular to the plane of the paper and so is the direction of the molecular chains.
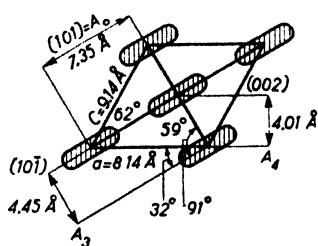
Fig. 89 Cross section through the elementary cell of cellulose II perpendicular to the fibre axis.

KRATKY[2] has shown that the crystalline regions occurring in cellulose fibres possess the form of long flat lamellae rather then that of rodlets. A diagrammatic picture of a crystallite with the fibre axis YY' vertical is reproduced in Fig. 90a. The lamellar plane corresponds to the A_o (101) plane in Fig. 89. Crystallographic planes perpendicular to the fibre axis of the crystallite, as e.g. the $II°$ (020) plane, are designated as diatropic, those

[1] M. POLANYI, Z. Physic, 7 (1921), 149.
[2] O. KRATKY, Z. physik. Chem., B 50 (1941) 255; Z. Elektrochem., 48 (1942) 587; cf. also W. A. SISSON, J. Phys. Chem., 40 (1940) 343; 44 (1940) 513.

parallel to this axis as paratropic planes (cf. Vol. I, Chap. I) *e. g.* A_o, A_3 en A_4 (corresponding to 101, 10$\bar{1}$ and 002 respectively). In the orientated fibre, the diatropic yield diffraction spots on the meridian M of the diagram (see fig. 91 showing a diagrammatic picture of the X-ray diffraction pattern of the principal planes). The interferences corresponding to the paratropic planes fall on the equator A of the X-ray photographs. The position of the planes A_3 and A_4 is indicated in Fig. 90b, representing a cross section through the lamellae.

If orientation is not perfect, there is a certain statistical distribution of angles between the crystal axes YY' and the fibre axis of the object. This statistical distribution, as well as the "average orientation" expressed in the orientation factor of the crystallites (cf. p. 590), may both be derived from X-ray analysis.

The difference between the diagrams for random and nearly perfect orientation have been shown in Fig. 83a and b; Fig. 92 shows a diagrammatic picture of the diagram of an intermediate case of partial uniaxial orientation. It is seen that the diffraction spots are sickle shaped lines, lying on circles around the point of incidence of the pencil of X-rays in the centre of the diagram. The intensity of the diffracted radiation, and hence, the intensity of blackening of the film along these circles, represents a direct measure of the distribution of the orientation of the corresponding crystallographic planes. This intensity distribution can be measured by photometric evaluation of the photographs.
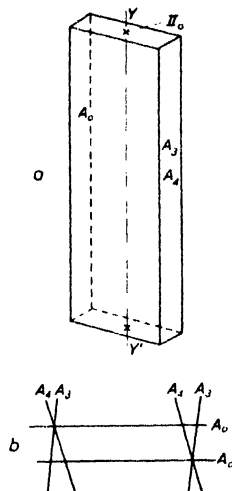
Fig. 90

Crystallite of cellulose II.

This procedure was first introduced by SISSON and CLARK [1], later perfected by KRATKY [2] and by the present writer *et al* [3]. It consist in determining the intensity of radiation along the circles as a function of the angular distance from the equator or from the meridian for the different interferences concerned. In as far as uniaxial fibre structure is concerned and merely the evaluation of average orientation is aimed at, the simplest procedure is to confine the measurements to the diatropic interferences. The intensity distribution along the diatropic sickles is — as mathematical analysis shows — a direct representation of the spatial distribution of the axes YY' of the crystallites (Fig. 90). If the intensity distribution as a function of the angular distance from the meridian be given by $I = F (\alpha)$, then the orientation factor f of the crystallites may be computed from the equation:
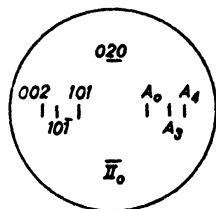
Fig. 91.

Diagrammatic picture of the X-ray pattern of orientated cellulose II (fibre axis vertical).

$$f = 1 - \frac{3}{2} \frac{\int_o^{1/2\pi} F(\alpha)\sin^3\alpha \ d\alpha}{\int_o^{1/2\pi} F(\alpha)\sin\alpha \ d\alpha} \tag{53}$$

[1] W. A. SISSON and G. CLARK, *Ind. Eng. Chem. Anal. Ed.*, 5 (1933) 296.

[2] P. H. HERMANS, O. KRATKY, and P. PLATZEK, *Kolloid-Z.*, 86 (1939) 245; P. H. HERMANS, O. KRATKY and R. TREER, *ibid.*, 96 (1941) 30.

[3] P. H. HERMANS, *Contribution to the physics of cellulose fibres*, Amsterdam-New York 1946; the same with J. J. HERMANS, D. VERMAAS and A. WEIDINGER, *Rec. trav. Chim.*, 65 (1946), 427.

A graphical procedure enabling one to derive the value of f from the photometer curves has been given by De Booys and Hermans[1].

The orientation factor has exactly the same meaning as that explained in § 8b. 3 (p. 590) and equation (53) may be written:

$$f = 1 - \frac{3}{2} \overline{\sin^2\alpha} \tag{54}$$

In practice, however, the diatropic interferences of cellulose fibres are usually too weak to allow of accurate results, and then the evaluation must be carried out starting from the much more intense paratropic interferences. This involves a somewhat more laborious procedure, which has been analyzed and described in detail[2].
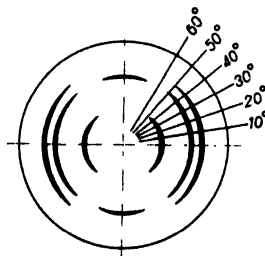
Fig. 92. Sketch of X-ray pattern of partly oriented cellulose

From the square sine of the average angles of orientation α_o and α_3 of the paratropic A_o and A_3 interferences respectively, the orientation factor of the crystallites can be calculated with the aid of the equation

$$f = 1 - \frac{3}{2} \overline{(\sin^2\alpha_o + \sin^2\alpha_3)} \tag{55}$$

It may be of interest to compare the orientation factors of the crystallites so obtained with the orientation factor computed from optical measurements (cf. Section 8 b . 3) which stands for the average orientation of the total fibre substance. In native ramie the two were found to be equal; in a series of regenerated fibres of increasing orientation, however, the optical orientation factor was found to be considerably lower than that computed from X-ray data at low degrees of orientation. At high degrees of orientation the optical and the X-ray orientation factor approach the same value. This indicates that in the process of orientation due to the stretching of the fibres in the swollen condition, the orientation of the crystallites runs ahead of that of the amorphous portion, a phenomenon which was also found in rubber, though to a much more marked degree. From data on the course of the two orientation factors some information on the percentage of crystalline substance in the fibres can also be derived[3].

Kratky[4] was the first to direct attention to the fact that sometimes a characteristic difference exists between the course of the orientation of the two paratropic planes A_o and A_3 which reveals itself as a difference in the values of $sin^2\alpha_o$ and $sin^2\alpha_3$ obtained. This observation is of importance in connection with the mechanism of orientation[5].

Besides data on the average orientation, the X-ray evaluation simultaneously yields information on the *distribution* of the orientation of the various crystallographic planes, which is, by itself, also of interest with regard to theoretical considerations concerning deformation mechanism.

Future research in this field will no doubt have to pursue the lines briefly explained here.

d Anisotropy of swelling

Isotropic gels, when allowed to swell or contract freely, change their dimensions to the same proportion in all directions. Orientated gels, on the contrary, show aniso-

[1] J. De Booys and P. H. Hermans, *Kolloid-Z.*, 97 (1941), 229.
[2] References see footnote 3 on previous page.
[3] See ref. 3 on p. 623.
[4] K. Baule, O. Kratky, and R. Treer, *Z. physik. Chem.*, B 50 (1941) 255.
[5] P. H. Hermans, J. J. Hermans, D. Vermaas, and A. Weidinger, *J. Polymer Sci.*, 1 (1946), 389, 393.

tropy of swelling. The phenomenon of anisotropical swelling is very common in natural objects like cell-walls, membranes, muscle fibres and so on, and has often been studied by botanists and zoologists. Several movements of animal and vegetable organs have been explained as being due to anisotropical swelling.

In uniaxially orientated objects, swelling is different in the direction of the axis of orientation and in the directions perpendicular to it. A spherical object will become a biaxial ellipsoid upon swelling or shrinking. In biaxially orientated objects swelling is different in three directions and a spherical object will become a triaxial ellipsoid. In the following we shall confine ourselves to the simplest case of uniaxial orientation.

Anisotropy of swelling can already be explained on the basis of the ancient micellar theory of VON NÄGELI [1]. Let us consider a fibre with uniaxial orientation consisting of rodlet shaped particles (Fig. 93). If the fibre is allowed to swell, and equal layers of a liquid are inserted between the particles, the change in diameter will surpass the change in length, since in the direction perpendicular to the fibre axis more intersticies per unit length occur than in the direction parallel to the fibre axis.

It will be clear that gels with an anisotropic net work structure will also exhibit anisotropy of swelling, though it is not so easy then, to give an exact description of the mechanism of swelling. As a matter of fact, the theoretical treatment of the case involves great difficulties and a satisfactory quantitative theory of the anisotropy of swelling in gels containing a molecular network remains to be developed.

One of the first attempts to give a theory on the anisotropy of swelling and to utilize this phenomenon to compute the degree orientation was made by HERMANS and DE LEEUW [2] in investigations on regenerated cellulose filaments. Swollen isotropic cellulose filaments, when subjected to a certain stretch, upon drying show a relative contraction in diameter, which is greater than the relative contraction in length. The anisotropy of swelling Q was defined as follows:

$$Q = \frac{\triangle d/d}{\triangle l/l}$$

where d and l are the diameter and the length of the filament and $\triangle d$ and $\triangle l$ the relative changes of these dimensions upon contraction. The theoretical relationship between Q and the special distribution of rodlet-shaped particles was derived. The particles were assumed to swell or contract solely in diameter and not in length. Though Q no doubt represents a measure of orientation, the quantitative relationship derived appeared to be not in agreement with the results of optical and X-ray investigations. The mechanism of swelling assumed here is doubtless a too simple one. Another theory was later given by J. J. HERMANS [3]. It appeared to be consistent with experiments in certain respects, without however satisfactorily solving the problem of relating the anisotropy of swelling with other measures of orientation. We shall revert to this matter in § 9.

Fig. 93. The explanation of anisotropical swelling on basis of the ancient micellar theory.

In cellulose fibres, it has been established, that Q changes from 1 in isotropic fibres to very high values up to ∞ 40 in highly orientated ones.

[1] A. FREY, Naturwiss., 15 (1927), 761.

[2] P. H. HERMANS and A. J. DE LEEUW, Naturwiss., 25 (1937) 524; Kolloid-Z., 81 (1937) 143, 300; 82 (1938) 58; cf. P. H. HERMANS and P. PLATZEK, ibid., 87 (1939), 296; 88 (1939) 68; 89, 344, 350.

[3] J. J. HERMANS, Rec. trav. chim., 65 (1946), 121

In the cases considered here, longitudinal and lateral swelling have the same sign. It may also occur, however, that the change of longitudinal and lateral dimensions is of opposite sign. This phenomenon can be frequently observed in animal hairs, tendons and many vegetable fibres upon swelling in a medium of sufficiently strong swelling power. Swelling is then accompanied by a lengthwise contraction and a lateral expansion. The latter surpasses the former to such degree that the total volume change is positive. This phenomenon was called "swelling-retraction". It was more closely examined for cellulose fibres, and a plausible explanation has been offered (cf. p. 647)[1]. According to this explanation, the underlying mechanism is intrinsically the same as that responsible for the thermo-retraction of cold stretched plastics and cold stretched raw rubber upon heating. We shall revert to these items in some more detail in § 9.

According to the above mentioned investigations, swelling retraction only occurs, if the degree to which the object is allowed to swell surpasses the degree of swelling which the object showed while it was previously orientated by stretching. Plastics exhibiting the phenomenon of thermoretraction also show swelling retraction, when treated with a swelling agent and, generally, there is a close analogy between the influence of swelling and that of a rise in temperature on the mechanical behaviour of macromolecular systems § 9.

Some theories have been put forward relating muscle contraction with swelling retraction but it seems rather doubtful whether the assumption of so a simple analogy is justified.

e   *Anisotropical mechanical behaviour*

Ftg. 94; Stress strain diagrams of cellophane iaken parallel and perpendicular to the direction of preferred orientation.

It is a well known fact that orientated macromolecular systems exhibit different mechanical properties in different directions. As far as chain molecules are concerned, the general rule that the tensile strength is greatest in the direction of preferred orientation of the chains, is conceivable. Fig. 94 shows the stress strain diagrams of a sheet of cellophane taken by loading it parallel and perpendicular to the direction of the greatest refractive index[2]. In the manufacture of the sheet a preferred orientation of the chains in the former direction has occurred. In orientated fibres the tensile strength of the material is maximum in the direction of the fibre axis. In highly oriented systems it is often observed that cleavage readily occurs in the direction of the orientation, indicating that the cohesive forces in the direction perpendicular to that of orientation are relatively weak. Fig. 95a and b are micrographs of the fragments obtained by pulverizing rigorously dried isotropic

[1] Cf. P. H. HERMANS, *Cellulosechemie*, 19 (1941) 117.
[2] G. VAN ITERSON, *Chem. Weekblad*, 30 (1933), 2.

and highly orientated cellulose filaments in a mortar. The isotropic objects yield irregular fragments similar to those obtained when pulverising glass or isotropic resins, whereas the orientated ones yield fibrous fragments[1]. A similar result is obtained when isotropic and extended rubber frozen in liquid air are pulverised[2].

a   b

Fig. 95.
Fragments obtained by pulverising well dried isotropic (a) and highly orientated (b) cellulose filaments.

§ 9. THE DEFORMATION OF GELS

a. Introduction

In a study on the mechanism of deformation of metals, GAUGH and WOOD[3] wrote: "The properties of the solid state are baffling in the extreme". As a matter of fact, the deformation of solid bodies represents a very difficult chapter of applied physics. It is well known that all natural phenomena in which the effects of forces acting between the atoms and molecules interfere are very difficult to deal with quantitatively. Even the theory of real gases which show a departure from the "ideal" behaviour has not yet been completely mastered. Is it then surprising that the theory of the liquid and particularly that of the solid state is entwined with a great deal of obstacles and difficulties?

The deformation of solid macromolecular systems and gels is, however, of paramount practical importance and, in the last decades, considerable progress has been made in the phenomenological study of the elasto-plastic behaviour of these objects, and also some more light has been thrown on the underlying mechanism of deformation in a few cases which may serve as a starting point for further reseach.

Whereas in the deformation of metals the "plasticity" of the crystals (internal gliding along certain preferred lattice planes) plays an important part, it has been

[1] P. H. HERMANS, Cellulosechemie, 18 (1940) 275.
[2] L. HOCK, Kolloid Z., 35 (1926) 40; Kautschuk, 3 (1927) 126.
[3] H. J. GAUGH and J. A. WOOD, J. Inst. Civ. Eng., 1937/1938, 249.

recognized that quite other principles govern the deformation of solid macromolecular systems. In the theoretical interpretation of the mechanism of deformation of macro-molecular systems, three materials have played an outstanding part in the develop-ment: rubber, cellulose and proteins. In the former case the theoretical results are most advanced. This is due to the particular nature of polymeric hydrocarbons in that their intermolecular cohesive forces are relatively weak and many changes of state may then be treated without taking energetic factors into account. The problem then, resolves itself into a purely geometrical one, involving configurational changes only.

The fact that deformation is intimately connected with the deformation of randomly kinked molecules first arose from the study of rubber as a mere hypothesis, which, in the meantime, may be considered as being very well supported.

X-ray investigation entered the field of rubber research, when it was found that beyond certain degrees of extension, where the chains are more and more parallelized in one direction, crystallisation phenomena occur, whereby new and more extended junction points are formed and a crystalline X-ray spectrum is observed. A charac-teristic feature in this class of substances is, that, according to the study of the X-ray diffraction pattern, the crystallites are very well oriented in the direction of extension from the very beginning of their appearance[1] (cf. p. 621). Upon continued extension they increase in number and size, but do not show any marked change in orientation. At maximum extension a typical fibre structure is formed, which can be stabilized by cooling. The structure of objects so obtained resembles that of well orientated cellulose fibres. The study of deformation in macromolecular systems with polar groups, like proteins and cellulose, where strong cohesive forces play a part, is facilitated when the swollen state is considered. Swelling diminishes intermolecular cohesion, since the forces acting are, partly at least, neutralized (screened, so to say), in that solvent molecules are bound. Such swollen systems exhibit rubberlike properties to some extent. Since, in the state of xerogels, these substances form resinous and brittle masses, which are hardly deformable at all, they have always been investigated in a more or less swollen condition and it may be expected that the problems involved will be the easier the higher the degree of swelling is chosen.

An intermediate position is taken by the polyesters and polyamides. Owing to the paraffinic nature of relatively large sections of the chain molecules concerned, they are less brittle in the xerogel state, though also here, the absorption of small quantities of a polar solvent like water, facilitates their deformation.

In an earlier section we have seen that the change of the X-ray diffraction pattern of certain proteins upon deformation, has given rise to the hypotheses of accom-panying configurational changes of molecular chains. According to this hypothesis, a regular pattern of chains folded up in a certain manner changes into another regular pattern of more straightened out chains. The development of the deformation theories, hence, emerged here from changes in the inner structure of the crystalline portion itself. The rôle played by the amorphous portion, which doubtlessly also occurs in these objects, has, thus far, not attracted much attention in this class of substances. The configurational changes of the shape of the chains, which is of a statistical nature in rubber-like substances, here follows a more definite path. On p. 606 we have seen that similar phenomena have been observed in synthetic polyamides.

---

[1] E. HEUSER and H. MARK, *Kolloidchem. Beih.*, 22 (1926), 63.

In the investigations on the deformation of cellulose, attention has, in first instance, been focussed on the accompanying change of the crystallite *orientation*. Among the three classes of systems mentioned, cellulose yielded by far the best X-ray diffraction patterns, over the entire range of states of orientation, offering the opportunity, not only of elucidating the structure of the crystalline lattice, but also to studying more accurately the change of the statistical orientation of the crystallites in the course of the deformation. Later it was recognized that the amorphous portion of cellulosic objects must also play a vital part in the phenomena. The idea of a netlike structure consisting of more or less convolute molecular chains linked up by crystalline junction points was developed and the theories of deformation accordingly modified, making use of concepts borrowed from rubber research. The crystallites themselves do not show any detectable change in the deformation process; it is only their orientation which changes. Also the quantity of the crystalline portion seems to remain sensibly constant under usual conditions.

Both in rubber and cellulose, and likewise in other linear polymers, the conception of a network structure now dominates as the general trend of all theoretical pictures.

Rubber is considered as a network of molecular entropy springs (cf. p. 568) with junction points consisting of chemical cross links or molecular entanglements. The principal theoretical viewpoints dealing with the extension of such a network structure have been treated of in Chapter IV, p. 123.

If rubber is investigated under conditions giving rise to increased intermolecular cohesion (e. g. if it partly crystallises on cooling or at high extensions) the theoretical treatment becomes more complicated and extremely difficult since then, besides entropy relations, energetic factors also come into the picture. The mechanical behaviour then also becomes more and more analogous to that of fibrous proteins and cellulose where intermolecular cohesion plays an intrinsic part.

Conversely, in the behaviour of a substance like cellulose upon deformation certain features remain akin to the general picture of network structure developed in rubber.

The crystallites act as the junction points of the network, the amorphous portion behaves more or less rubberlike, but the flexibility of the chains, the degree of convolution and further, the average number of statistical chain elements between the junction points, seem to be different, all these quantities being smaller than in rubber. Furthermore, energetic factors, bound up with the much greater intermolecular cohesive forces in cellulose, change the picture and render it far more complicated.

To deal with the deformation and the so called mechanical properties of macromolecular systems in some degree of thoroughness and completeness would lead us too far afield in this chapter. We shall therefore confine ourselves to the discussion of some selected topics which may convey a general idea of the current views with respect to the subject. We shall first endeavour to give a separate concise treatment of the cases of rubber and cellulose and then let follow a section dealing with a comparative discussion of the mechanical behaviour of these two systems, which, more or less, represent extreme cases. The behaviour of other systems may be regarded as being more or less intermediate between these two.

All these objects claim a certain practical importance; they exhibit a high extensibility and a considerable permanent coherence and mechanical strength. Many other gels, like those of agar agar, starch and gelatine show a peculiar deform-

ability and elastic behaviour. They, however, do not endure large extensions without breaking. The coherence of the gel frame obviously does not stand the stresses required to bring about a far going change in orientation. Since very little fundamental information is available referring to these cases, we shall not further occupy ourselves with them here. It should be noted, that very diluted cellulose gels and swollen rubber also belong to this class.

b. The deformation of rubber

b. 1. *General picture of the deformation mechanism of rubber*

The principles of the modern physical theories of the deformation of rubber xerogels have already been dealt with in Chapter IV, p. 123. If a strip of raw rubber is rapidly extended, the molecules, which initially assumed randomly kinked forms, are stretched too. The more extended shape of the chains is a statistically less probable one corresponding to a lower entropy. When the piece of rubber is rapidly released again, it reassumes its original form in that the chains return to their most probable configurations. The entropy-character of rubber elasticity has been proven in that it exhibits a positive temperature coefficient.

The internal friction in rubber is low, since the intermolecular cohesion is weak and the adjacent chains may relatively readily glide or rotate along each other. "Micro-flow" is easy in rubber. This is also the reason that, in raw unvulcanized rubber, besides microflow, "macro-flow" is observed too, especially at higher temperatures. If the extension is carried out very slowly or if the rubber is held in the extended state for a longer time, the molecules may also glide lengthwise along each other. On release of the stress it is then observed that the piece of rubber does not reassume its initial dimensions but shows a permanent elongation, whose magnitude is dependent on the duration and the temperature of the experiment. This is essentially the case represented by the mechanical models having a piston and a spring in parallel, reproduced in Vol. I, Chapter I. Though a permanent extension is reached, the rubber is still isotropic after its recovery.

In vulcanized rubber the molecular chains are cross linked by chemical bonds[1], which inhibit macro-flow. That is why vulcanized stocks are more perfectly elastic than raw rubber (cf. the models having piston and spring in series.) The chemical cross links, randomly distributed over the mass of rubber, act as permanent junction points. A netlike structure is obviously present in vulcanized rubber.

In rubber there is, however, still another factor which plays a very important rôle in preventing lengthwise gliding. This is the phenomenon of *crystallisation*, which occurs beyond a certain degree of extension (cf. also p. 621)[2]. The molecular chains pulled into a better parallel alignment crystallise in several favourable regions. Heat of crystallization is given out and the chains are held together by the lattice energy. It would seem that these centres of crystallisation correspond to the regions where the internal tensions are maximum. This is illustrated by the fact, already mentioned before, that the crystallites appearing are always nearly perfectly oriented in the direction of stretch. The molecular fringes protruding from the crystallites

---

[1] Some authors believe that other factors may play a part in vulcanization as e.g. the introduction of strongly polar groups or substituents of larger size, which inhibit the mutual lengthwise gliding of the chains. See *e.g.* the survey by W. F. BUSSE, *Ind. Eng. Chem.*, 31 (1939), 1391.

[2] *e.g.* L. R. G. TRELOAR, *Trans. Faraday Soc.*, 36 (1940), 538; 37 (1941), 84.

have a more random orientation. CLARK [1] has shown that a piece of extended vulcanized rubber, which exhibited a perfectly orientated X-ray fibre pattern, when cooled to a low temperature while maintaining the extended state, after some time exhibits an other X-ray pattern with sickle shaped interferences quite analogous to those given by partially oriented cellulose fibres, where orientation has resulted from pulling crystallites already existent from a random array into a preferred orientation. The freezing obviously causes a further crystallisation and the new crystallites now show a less perfect orientation, indicating that the amorphous portion from which they were formed was less well orientated.

Rubber thus frozen in the extended state does not recover elastically upon release of stress; it has a "permanent set". In other words, retraction is blocked by cohesive forces. On warming up, however, retraction occurs at once. Under ordinary conditions, at room temperature, however, the lattice forces of the crystallites formed upon extension are not strong enough to prevent the rubber from retraction. The entropy factor then surpasses the work of crystallisation.

The *blocking of retraction* of rubber in the frozen condition can also be removed by swelling. This will be clear, since the swelling agent interferes with the cohesive forces.

Rubber stocks showing optimal tensile strength are generally those which show the highest degree of crystallisation [2]. This is why swelling has a detrimental effect on tensile strength.

It would seem that the technical importance of vulcanization is (amongst others) to prevent lengthwise permanent flow until extensions are reached where crystallization sets in, which now further takes over this function to a still more effective degree [3]. BUSSE has described very interesting experiments involving the thorough vulcanisation of lightly vulcanised rubber in the extended condition by subjecting the stretched samples to the action of sulphur chloride vapours at room temperature [4]. He so obtained filaments showing permanent set and possessing an unusually high tensile strength of 12 kg/mm² (which is comparable to the strength or ordinary viscose rayon). In the non vulcanised condition the tensile strength of the same material was not more than about 0.01 of this amount. The crystalline state has, so to say, been fixed, by a thorough interlinking of the chains, thus inhibiting recovery to randomly kinked forms.

b. 2. Deformation of frozen rubber, gutta percha and balata

Let us now see what happens if isotropic rubber is frozen. Crystallisation occurs, but the crystallites are now randomly orientated. This can be most readily demonstrated with raw unvulcanised rubber [5]. Rubber thus frozen is hard and stiff and has lost its elastic character. With regard to a general understanding of the mechanical

[1] G. L. CLARK, *Ind. Eng. Chem.*, 31 (1939), 1397; cf. also the experiments of VAN DER WYCK described in K. H. MEYER, *Die Hochpolymeren Verbindungen*, Leipzig 1940, p. 121.

[2] J. E. FIELD, *J. Appl. Physics*, 12 (1941), 23.

[3] For a survey of the effects of vulcanization in rubber, cf. H. L. FISHER, *Ind. Eng. Chem.*, 31 (1939), 1381; W. W. VOGT, *ibid.*, 31 (1939), 1385; W. F. BUSSE, *ibid.*, 31 (1939), 1391; G. L. CLARK *loc. cit.*

[4] W. F. BUSSE, *J. Phys. Chem.*, 36 (1932), 2862.

[5] A. VAN ROSSEM and J. LOTICHIUS, *Kautschuk*, 5 (1920), 2; W. H. SMITH and C. P. SAYLOR, *J. Research, Nat. Bur. of Stand*, 21 (1938), 257.

properties of xerogels, it is of considerable interest to compare the mechanical behaviour of rubber thus frozen with that of unfrozen rubber. We may then also direct attention to the related rubber-like materials gutta percha and balata. These are natural high polymeric hydrocarbons of identical chemical constitution as rubber, except the spatial configuration at the double bonds (cf. Chapter II), rubber having the cis- and the others the trans-configuration. Gutta percha and balata are thereby characterized that the melding points of their crystalline modification lies beyond room temperature. They therefore show rubber-like elasticity only at higher temperatures. We shall revert to this subject in section 9 d, dealing with stress-strain relations in general. We shall here confine ourselves to stating that, upon extension of the isotropic crystalline objects, the molecules must be torn loose from their lattice arrangements in order to allow their orientation in the direction of extension. Existing randomly oriented crystalline regions are destroyed and other oriented ones rebuilt in this process. There are many indications that processes of this kind actually often occur in the deformation of macromolecular systems.

b. 3. *Optical and X-ray behaviour*

Owing to the overall orientation of the molecules, rubberlike substances become birefringent upon extension. The double refraction rises approximately proportional to the degree of extension provided crystallisation phenomena do not interfere.

Some very interesting studies on the optical behaviour of rubber in connection with crystallisation phenomena have been published by THIESSEN and WITTSTADT[1] and by TRELOAR[2] and are in conformity with X-ray data. It was observed that the birefringence of stretched samples, when held at constant elongation, continued to rise for a considerable time and this rise was ascribed to crystallisation. The phenomenon only occurred if the initial extension surpassed a certain percentage. Below this degree of extension which, according to X-ray investigations, coincides with the limit beyond which crystallisation sets in, the birefringence remained constant. This limit corresponds to a higher degree of extension, the higher the temperature of the experiment is chosen. TRELOAR, who also carried out experiments at low temperatures (0°) with raw rubber, conditions whereby crystallisation is considerably favoured, found a striding parallelism between the changes of birefringence and those of the density of the samples and a close correlation with the plastic flow and elastic recovery phenomena. The birefringence of a rubber sample in a given condition depends on three factors: a) the degree of crystallisation, b) the orientation of the crystallites, c) the orientation of the amorphous portion (compare p. 593).

In a particular experiment of TRELOAR, a piece of rubber which has been extended to 260% elongation, and allowed sufficient time at 0° for the complete recrystallisation, was subjected to a further deformation at this temperature, whereby its elongation was increased to 470% (compared with the unstretched length). In spite of the increased elongation, the birefringence immediately fell, and it was only after the lapse of time that it eventually rose to a value higher than that which it possessed at the original elongation. This clearly demonstrates that, upon the second deformation, the crystallites originally present were torn assunder to a certain extent, this

[1] P. A. THIESSEN and W. WITTSTADT, *Z. physik. Chem.*, B 29 (1925) 359.
[2] L. R. G. TRELOAR, *Trans. Faraday Soc.*, 37 (1941) 84; 43 (1947), 84.

being required in order that the molecules may move into the new positions which the extension compels them to take up.

It could be also concluded from the optical data that, even at very low extensions, the degree of orientation of the crystalline nuclei is very much higher than that of the rubber molecules as a whole, a conclusion also arrived at by X-ray research.

Attempts to compute the amount of crystallisation in rubber from X-ray data have been published by FIELD and GOPPEL[1]. The X-ray diagram of isotropic amorphous rubber shows a diffuse ring, the intensity of which decreases upon extention of the material according as the intensity of the diffraction spots due to the crystalline portion is enhanced. In measuring the intensities as a function of stretch, the percentage of cristallinity could be computed. Though, as GOPPEL showed, FIELD's figures were too high, the majority of his conclusions remain unaffected. He could establish a close correlation between crystallisation, permanent flow and mechanical properties, and clarify a number of problems as to the influence of small variations in the composition of rubber stocks on the mechanical properties of the vulcanisate. According to GOPPEL the maximum percentage of crystalline substance is in the order of 40 %.

b. 4. Quantitative theories of rubber deformation

Theoretical attempts to cast the deformation of rubber-like systems into a mathematical form have been given by KUHN and GRÜN[2], HERMANS[3], TRELOAR[4], FLORY, and REHNER[5]. All these theories are fundamentally based on the statistics of the chain molecule as first developed by KUHN[6]. (Cf. Chapter IV, p. 93 by J. J. HERMANS). The picture employed is essentially that of a network structure consisting of a system of uniformly distributed junction points interconnected by flexible molecular chains, which tend to assume the most probable configuration throughout all conditions. The extensibility v of such a structure is of course limited, though it may be very considerably depending on the degree of convolution of the chains in the original condition and the number of junction points per unit volume. Its maximum value may be computed to be

$$v_{max} = 2.2 \sqrt{N} \qquad (56)$$

where N is the average number of "statistical chain segments" (cf. Chapter IV) occurring between the junction points.

From this equation we can see that $N = 36$ will yield a value of v_{max} in the order of magnitude of 10, which corresponds to roughly the maximum relative elongation shown by rubber. For a given number of atomic links per chain, N will be the smaller the less flexible is the chain. Stiffer molecules will hence give rise to a lower extensibility. If the number of junction points per unit volume increases,

[1] J. E. FIELD, *J. Appl. Phys*, 12 (1941) 23; J. M. GOPPEL, *Appl. Sci. Res.*, 1 (1947), 3.

[2] W. KUHN and F. GRÜN, *Kolloid-Z.*, 101 (1942) 248.

[3] J. J. HERMANS, *Kolloid-Z.*, 103 (1943) 210.

[4] L. R. G. TRELOAR, *Trans. Faraday Soc.*, 39 (1943) 36, 241; 43 (1947). 278.

[5] J. J. FLORY and J. REHNER, *J. chem. Phys.*, 11 (1943), 512, 521; J. J. FLORY, *Chem. Rev.*, 35 (1944) 51.

[6] W. KUHN, *Kolloid-Z.*, 68 (1934) 2; 76 (1936) 258; 87 (1939) 3; *Z. angew. Chem;*, 49 (1936) 858; 51 (1938) 640; 52 (1939) 289; cf. also the papers of F. H. MÜLLER, *Kolloid-Z.*, 95 (1941) 181, 307.

N, the number of chain elements per chain will, of course, decrease giving also rise to a lower extensibility.

The mathematical problem of deformation now consists in calculating how the network structure concerned behaves upon extension in one direction. KUHN and GRÜN as well as FLORY and TRELOAR base themselves on the principle of "affine deformation", i. e. they assume that the distances between the junction points change in the same proportion as the macroscopic dimensions of the system (cf. Chapter IV). If l and l_o be the length of the body in the direction of elongation after and before the extension respectively, then $v = l/l_o$ is the *degree of extension* and $\gamma = v - 1$ the *specific elongation*. If the volume of the body remains constant, its relative lateral contraction will be $v^{-\frac{1}{2}}$. The entropy change as a result of stretching, the optical anisotropy of the system and the retractive force can then be calculated. The work of all the various authors mentioned lead to essentially the same equation.

The anisotropy of polarisability of the extended body is given by the expression

$$\varLambda = \frac{1}{5} \; G_o \; (a_1 - a_2) \; (v^2 - v^{-1}) \tag{57}$$

where $(a_1 - a_2)$ is the anisotropy of the statistical chain element[1]. Since

$$v^2 - v^{-1} = 3 \, (v - 1) + (v - 1)^3 - (v - 1)^4 + (v - 1)^5 - \ldots \text{ etc.} \tag{58}$$

equation (57) may be thus simplified for small values of v

$$\varLambda = \frac{3}{5} \; G_o \; (a_1 - a_2) \; (v - 1) \tag{59}$$

Expressing the anisotropy in terms of birefringence instead of polarisabilities it can easily be shown[2] that equation (57) becomes:

$$n_{,,} - n_{\perp} = k \; \frac{(n^2 + 2)^2}{n} \; (a_1 - a_2) \; G_o \; (v^2 - v^{-1}) \tag{60}$$

where k is a numerical constant, $n = \frac{1}{2}(n_{,,} + n_{\perp})$ and G_o the number of chains per unit volume (cf. p. 97)[3].

The retractive force (stress with reference to the initial cross section) was calculated by KUHN to be

$$K = c \; G_o \; (v^2 - v^{-1}) \tag{61}$$

The value of K will be the higher, the larger G_o and hence, the greater the number of junction points per unit volume. From equations (60) and (61) follows that the retractive force and the birefringence should be proportional. This is actually borne out by experiments in the initial phases of extension up to not too large extensions[4].

These theories were invariably based on the assumption that the number of statistical chain elements between the junction points is large. Calculations for a

---

[1] Cf. KUHN and GRÜN, *loc. cit.*

[2] In the calculation it is assumed that $n_{,,} - n_{\perp}$ is small as compared to $n_{,,}$ and $n_{\perp}$, which always holds in the systems considered.

[3] J. J. HERMANS endeavoured to free the calculation from the arbitrary assumption of affine deformation of the junction points and came to more complicated formulae, which for *small* elongations yield an identical result apart from a slightly higher value of k.

[4] Cf. F. H. MÜLLER, *Kolloid-Z.*, 95 (1941) 181, 307.

restricted number of statistical chain elements ("short chains") have been carried
out by HERMANS[1] and lead to interesting results. The stress strain diagrams of rubber,
when plotted against the force of extension calculated on the original cross section
of the sample are slightly S-shaped. Whereas the earlier quantitative theories only
succeeded in explaining the initial part of these curves[2], the "short chain theory"
is capable of accounting quantitatively for the shape of the curve over a considerable
range of extensions. Furthermore, the short chain theory would seem to be of particular
importance for cellulose gels where the number of chain elements N per chain seems
to be very low (cf. p. 646).

A particular feature of the short chain theory is also that it does not lead to a
proportionality between birefringence and retractive force. In this theory their
relationship depends on the value of N.

b. 5. Relationship between deformation and swelling

A very interesting development of the theory was initiated by FLORY in connecting
the problem of swelling with that of deformation. It has been argued on p. 571 that
it is inevitable to connect the contraction of swollen gels upon drying with a folding
or crumpling up of the molecular chains between the junction points. Conversely,
the swelling of network structures of the kind discussed now should be visualized
as an isotropic expansion of the network, whereby the individual chains are stretched
in exactly the same way as upon deformation of the network[3]. Combining KUHNS
principles with the thermodynamical considerations on swelling and dissolution,
a relationship between the degree of swelling and the degree of cross linking can be
derived. The theory can now also be extended to the deformation of swollen objects[4].
If the average number N of "statistical chain elements" (in Kuhn's sense) is large
the theory requires an initial increase in volume on stretching[5]. This volume increase
is restricted to moderate elongations since it is clear that at more extreme degrees
of stretch the meshes of the network will eventually be closed, and the swelling
medium be pressed out. It depends on the value of N where this effect becomes
predominant; if N is sufficiently small the volume of the network will decrease from
the very beginning.

Experiments of this kind seem not to have been carried out as yet with rubber
gels. They are not easily realized since most swollen rubber-like substances break
at relatively low extensions. Since N is relatively large here, initial increase in volume
should be expected.

In swollen cellulose gels many observations have been reported, showing that
very considerable contractions in volume to even less than half the original volume
may occur on stretching[6]. (Cf. Fig. 96).

---

[1] J. J. HERMANS, *J. Colloid Sci.*, 1 (1946), 149 (cf. also Chapter IV, p. 126).
[2] Cf. *e. g.* L. R. G. TRELOAR, *Trans. Faraday Soc.*, 40 (1944) 109; P. J. FLORY, *Chem. Rev.*,
35 (1944) 51.
[3] The intimate relationship between swelling and deformation has since long been recognized
in investigations on cellulose gels by the present author (cf. next section).
[4] P. J. FLORY, *J. Chem. Phys.*, 11 (1943) 512, 521. (Also see Chapter IV, p. 128)
[5] Compare J. J. HERMANS, (*General Discussion on Swelling and Shrinking*, held by the *Fara-
day Soc.*, London 1946 (in press).
[6] *e. g.* P. H. HERMANS and A. J. DE LEEUW, *Kolloid-Z.*, 81 (1937) 300; P. H. HERMANS, *Cellulose-
chemie*, 19 (1941) 117.

It would lead us to far afield to enter into further details here. This brief survey may serve to convey some general ideal of the course of the development which the quantitative theoretical treatment of the concept of the network structure of gels has recently taken. Rubber-like systems are the objects "par excellence" for studying the problems and the principles concerned with the deformation of macromolecular systems, since they represent the simplest ones from a physical point of view.

c. The deformation of cellulose gels

c. 1. Introduction. The use of model filaments for studies on deformation of cellulose

For reasons already explained in the foregoing, the study of the deformation of cellulose gels is entwined with much greater theoretical difficulties. In order to minimize the latter as far as possible, two conditions should be fulfilled. Firstly attention should be focussed on swollen objects, thus diminishing the complications due to energetic molecular interaction, and secondly isotropic objects should be available as a starting material.

Fibres from natural scources as well as the commercial artificial objects all represent more or less orientated systems not fulfilling the latter condition. They all result from a previous process of orientation. The sole objects hitherto investigated fulfilling both conditions are the isotropic swollen filaments as introduced by HERMANS and DE LEEUW[1], and later also used by KRATKY[2] *et al.* These objects can be very easily prepared and handled. They may be obtained in various degrees of swelling up to $q = 16$ and over. They will be referred to as "model filaments" below. Their preparation essentially consists of allowing a cellulose xanthate solution (viscose) to run very slowly from a capillary tube into an ammonium sulphate solution (spec. gr. 1.08) without exerting any appreciable pull on the filament formed, which sinks to the bottom of the vessel by its own gravity. Cylindrical gel-filaments consisting of cellulose xanthate are thus formed[3], which are still soluble in water, but may be easily transformed into cellulose filaments by a treatment with dilute acid. The degree of swelling depends on the composition of the viscose and of the coagulating bath, but, under usual conditions, is around 10—15 for the xanthate filaments and about half as much for the cellulose filaments. Upon exposing them to the air, air-dry isotropic cellulose filaments are obtained which, upon re-swelling in water, attain a swelling degree of about 2.2 (cf. p. 640).

The extensibility of the swollen objects depends on the conditions of preparation and on the degree of swelling, but lies in the order of magnitude of 100% (for the air-dry filaments, which contain about 15% water, the extensibility is usually smaller and below 100%). Fully dried filaments cannot be extended at all, they are hard and brittle as glass and can be pulverized in a mortar. The presence of water is essential in order to make a cellulose gel deformable. Dry cellulose is comparable with rubber cooled to temperatures far below zero, say — 100°. In both cases the cohesion between the chains is so considerable that there is no more question of an internal mobility, which is essential in order to render large deformations possible.

---

[1] Cf. preceding footnote.
[2] See ref. 2, p. 641.
[3] The degree of substitution corresponds usually to about 3—4 xanthate groups per 10 glucose residues and is, hence, very small so that the objects can be almost considered as being cellulose gels.

Highly swollen cellulose filaments are very elastic. At extensions up to, say 30—50% they show a spontaneous recovery whereby the greater part of the elongation is redressed, provided the experiment be carried out rapidly. If kept in the elongated condition for a longer time or at higher extensions, the recovery is relatively small. As will be further argued later, the permanent extension is not due to a process of macro-flow as in unvulcanized rubber, but to *a blocking of the mechanism which should bring about recovery*. A lengthwise gliding of the chains along each other would be an extremely improbable process in cellulose when taking into consideration the strong intermolecular cohesion in the junction points.

The volume of the highly swollen gels is considerably reduced on deformation as demonstrated by Fig. 96. To use a rough picture, the network of the gel is, obviously, "squeezed out" by the deformation. A considerable amount of water is pressed out of the gel. As to the question why no recovery takes place on release of stress, we may assume that during the compression of the network new secondary junction points are being formed where the chains meet each other in favourable positions. This might essentially be the mechanism of the blocking.

The elasticity at low and moderate degrees of extension may be connected with a mechanism similar to that in rubber.

Fig. 96. Swelling degree q of swollen modelfilaments of cellulose in dependence on the degree of extension v. X. Xanthate filaments, F. freshly prepared cellulose filaments.

The chains extending themselves between the primary crystalline junction points (cf. p. 497) will, just as in rubber, possess a randomly kinked statistically most probable configuration and may be responsible for the elastic behaviour.

The formation of new secondary junction points by which the mechanism of recovery is blocked may be readily compared with the blocking which occurs when extended rubber is frozen. In rubber, the blocking can be released either by raising the temperature or by swelling. In cellulose only the latter factor comes into consideration.

If stretched model filaments showing a permanent extension are allowed to swell in a suitable swelling agent, considerable retractions at once occur. The filaments become shorter and simultaneously their volume increases. This phenomenon was termed swelling retraction. Though a complete reversal of the extension can be only reached under extreme and difficultly realizable conditions, the contraction, as far as it goes, actually represents a true partial reversal of the process of elongation[1].

As will be set forth later in more detail, there is, however, an intrinsic difference between the cellulose gels and rubber, in as far that cellulose chains are considerably stiffer than those of rubber. The length of the statistical chain element comprises about 20 glucose residues (cf. Ch. IV, p. 109). A cellulose chain of given length is

---

[1] Cf. P. H. HERMANS, *Cellulosechemie*, 19 (1941), 117.

therefore less convoluted in its most probable configuration than a rubber chain of equal length. Furthermore, there are certainly more junction points per unit volume in cellulose gels, hence, the chains comprise merely a few statistical sections in KUHN's sense. This is in conformity with the much smaller maximum extensibility of the cellulose gels as compared to rubber. Recalling the formula given on p. 568, N would even be of the order of magnitude of only 1 or 2.

c. 2. Deformation and orientation. Relationship with swelling

Isotropic model filaments after having been subjected to a permanent extension represent orientated systems. They show anisotropy of swelling, optical anisotropy and their X-ray diagram reveals a preferred orientation of the crystallites. At maximum extension a high degree of orientation may be reached. It is highest in xanthate filaments and lowest in filaments stretched in the air-dry state.

Discussing orientation in cellulose gels, one will always have to discriminate between that of the crystalline and that of the amorphous portion. A measure of the former may be derived from X-ray data. Orientation computed from birefringence represents a measure of the average orientation of both components superimposed on each other.

In Table 13 the values of the optical orientation factor f_o and that derived from X-ray data (cf. Sections 8 b 3, and 8 c 2, p. 623) are given for xanthate filaments ($q = 13$) and freshly prepared cellulose filaments ($q = 7.0$) at various degrees of previous stretching[1]. (The data refer to the objects dried in air after their preparation[2].

TABLE 13

OPTICAL ORIENTATION FACTOR f_o AND THAT DERIVED FROM X-RAY DATA f_x FOR CELLULOSE FILAMENTS
IN DEPENDENCE ON THE DEGREE OF PREVIOUS STRETCHING

	Percent stretched	f_o	f_x	f_o / f_x
xanthate filaments $q = 13$	25	0.03^7	0.06^3	0.59
	50	0.27^6	0.37	0.75
	75	0.61^1	0.73	0.84
	100	0.79^1	0.89^1	0.89
	125	0.90^4	0.90	1.0
fresh cellulose filaments $q = 7.0$	25	0.03^7	0.08^5	0.43^5
	50	0.18^3	0.25	0.73
	75	0.44	0.63^1	0 70
	100	0.64	0.78^3	0.82

It is seen from the table that, at lower degrees of stretch, the orientation of the crystallites (f_x) runs ahead of the overall orientation of the total fibre substance (f_o), a phenomenon analogous to that observed in rubber, though being less marked than in the latter case. We shall once more revert to this phenomenon later. For comparison

---

[1] P. H. HERMANS, Contribution to the Physics of Cellulose fibres, Amsterdam-New York, 1946, p. 186. It is recalled that the symbol q stands for degree of swelling.
[2] The birefringence of the fibres may be calculated from f_o, using the data given in Table 8 on p. 592.

TABLE 14

OPTICAL AND X-RAY ORIENTATION FACTORS OF SOME NATURAL AND RAYON FIBRES

Fibre specimen	f_o	f_x	f_o / f_x
native ramie	0.97	0.97	1
mercerized ramie (stretched)	0.98	0.98	1
viscose rayon A . . . . _.	0.62[5]	0.78	0.80
„ „ B	0.74[5]	0.89	0.84
„ „ C	0.88	0.91	0.98
Lilienfeld rayon	0.91	0.92[5]	0.98[5]
Cuprammonium rayon	0.74[5]	0.86	0.86[5]

we add some corresponding data referring to natural and commercial artificial cellulose fibres in Table 14. (Three species of viscose rayon spun with increasing degrees of stretch from A to C are included in the table).

The stress required to bring about extension at a given degree of orientation increases towards lower degrees of swelling and, of course, it increases also with the orientation. In the air-dry objects, as a result of the high internal friction, the tension to bring about orientation increases to such an extent that the structure cannot stand it beyond a certain degree of orientation and the filament consequently breaks before orientation has been carried to completion. High degrees of orientation cannot be reached by stretching air-dry filaments. Optimal conditions must be sought in order to reach a maximum degree of orientation. It is clear, however, that, owing to the irregularity of the network structure formed upon the coagulation of the gel, complete orientation will never be reached. Some meshes of the network will be pulled straight and tight together, before others have reached this state. Upon further extension, the former will break,

Fig. 97. Diagrammatic representation of the structural changes in a cellulose gel upon shrinking and upon deformation.

thus weakening the structure and giving rise to a subsequent rapid rending of the entire structure. (This picture is essentially consistent with the idea of flaw formation, which is generally recognized as governing the tensile strength of solid bodies).

A rough diagrammatic representation of the picture of deformation referred to in the foregoing is given in Fig. 97.

Let us assume that the degree of swelling of the primary isotropic gel be 8. At a_1 one single mesh out of the isotropic gel is shown, having a linear dimension in some arbitrary unit of say 200. The junction points are represented by dots. Upon

deformation, the configuration shown at a_2 will result. If the degree of extension is 2, the linear dimension in the direction of elongation will, hence, become 400. If the primary gel is allowed to contract in the isotropic condition by drying, its linear dimension will become $200 \cdot 8^{-1/3} = 100$. The kinkyness of the chains is thereby considerably enhanced (see Fig. 97b_1). If now the dry gel is extended to reach the same degree of orientation as before, its linear dimension will have to be increased from 100 to 400 (Fig. 97b_2). Hence, the dry gel requires a much greater extension to attain a given degree of orientation than the swollen one. This is exactly what has been qualitatively found. The rate of orientation as a function of elongation is smaller, the lower is the swelling degree of the gel. (This, however, only applies when exclusively comparing filaments originating from the same primary gel). HERMANS and PLATZEK[1] have shown that the extensibility of a given modelfilament is approximatively proportional to the cub. root of its degree of swelling. Assuming that the filaments
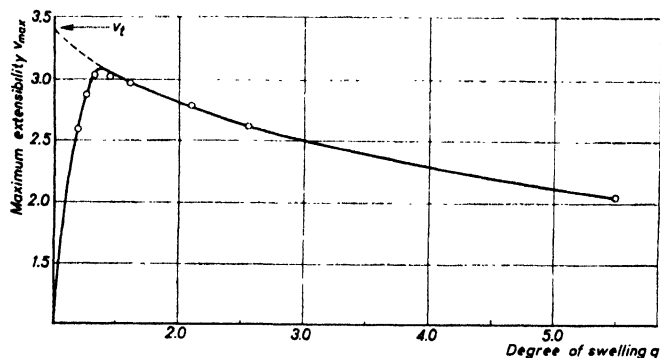
Fig. 98. Maximum extensibility on isotropic model filaments in dependence on the swelling degree q.

breaks at approximately the same degree of orientation (which is roughly true) this is exactly what should be expected from the diagram shown in Fig. 97. In Fig. 98 the experimental figures of maximum extensibility are plotted against the original degree of swelling of the isotropic filament (dots). The broken line is the curve calculated on basis of the equation: $v_{max} = v_t q^{-1/3}$, where v_t is the extrapolated value of the extensibility for $q = 1$. It is seen that the equation holds over a large range of q.

At low values of q, another phenomenon occurs and v_{max} drops to zero if the filament is completely dried. In this region the increased internal friction interferes and the filament breaks at lower and lower degrees of extension. The air dry condition, with about 15% moisture, corresponds to $q = 1.17$. Dried filaments after reswelling in water reach $q = 2,2$, a figure lying near the point of maximum extensibility. (Xanthate filaments fall outside this picture; their extensibility is always greater than corresponds to their degree of swelling on account of Fig. 98).

Model filaments are particularly suitable objects for quantitative studies on the course of orientation as a function of elongation and have been profitably used in several investigations of this kind[2]. However, before going further into this matter, we shall first briefly deal with the development of some of its theoretical aspects.

---

[1] P. H. HERMANS and P. PLATZEK, Kolloid-Z., 97 (1941), 329.
[2] A comprehensive survey of the subject, its theoretical aspects and the literature concerned will be given in the present authors book ,,Physics and Chemistry of Cellulose Fibres" (to appear soon).

a. Quantitative theories on cellulose deformation. The first who endeavoured to attack the problem of orientation in cellulose gels from a quantitative theoretical side was KRATKY[1]. In accordance with the theoretical ideas prevailing at the time, he merely focussed attention on the crystalline portion of the gel and considered the case of randomly orientated crystalline rodlets embedded in a viscous matrix which is subjected to an elongation at constant volume. As in almost any theory on the deformation of solids, the principle of "affine deformation" was used as a basis of calculation. It is assumed that the rodlets are simply dragged along with the matrix, and that any microscopic volume element of the latter follows the dimensional changes of the macroscopic body. Fig. 99 shows how three arbitrarily chosen rodlets occurring in the isotropic body (a) would change their position upon elongation (b) of the body according to the principle of affined deformation. For an initially random distribution of the rodlets in the isotropic state it can be calculated what their spatial distribution will be after deformation. The mathematical theory shows that orientation is only complete at infinite elongation and this seemed to be inconsistent with many observations on cellulose gels revealing that almost complete orientation can be reached at elongations of the order of 100 %.

Later KRATKY *et al*[2] drew attention to the fact that the very considerable contraction in volume occurring upon stretching highly swollen cellulose gels (Cf. Fig. 96) should be accounted for. According to the principle of affine deformation such a contraction in volume should give rise to a higher degree of orientation than an equal relative elongation at constant volume. If the initial and final degrees of swelling be q_i and q_f respectively and v the relative elongation, then it can be easily shown

Fig. 99. Diagrammatic representation of the principle of affine deformation.

that the same orientation would be attained as at a relative elongation $v(q_i/q_f)^{1/3}$ at constant volume. This was called by KRATKY the equivalent degree of elongation v_a.

It is important to note that, according to this theory, changes in degree of swelling may influence orientation. This is, of course, not the case if the gel is isotropic. If, for instance, the isotropic body in Fig. 99a shrinks isotropically to the body in Fig. 99c no change of orientation will occur. On the other hand, orientated gels always exhibit anisotropic swelling. If the anisotropic gel represented in Fig. 99b is allowed to shrink, its relative change in diameter will be greater than that in length (Fig. 99d) and, if the principle of affine deformation also holds for the process of shrinking, a considerable increase in degree of orientation should take place.

Recently HERMANS *et al*[3] have shown by quantitative X-ray analysis of cellulose filaments (stretched in the swollen state to various extents) that very little if any change in orientation occurs on drying, thus proving conclusively that the process

[1] ECKLING and O. KRATKY, *Naturwiss.* 18 (1931), 461; O. KRATKY, *Kolloid-Z.*, 64 (1933), 213
[2] B. BAULE, O. KRATKY, and R. TREER, *Z. physik. Chem.*, B 50 (1941) 255.
[3] P. H. HERMANS, J. J. HERMANS, D. VERMAAS and A. WEIDINGER, *J. Polymer Sci.*, 2, (1947), 632.

of drying can even not approximately be accounted for by the principle of affine deformation.

There remains the question of whether this principle can nevertheless account for the process of orientation upon stretching in the swollen state. KRATKY *et al*[1] have claimed that it can in first approximation. Postulating crystallites having the form of lamellae instead of that of rotatory symetrical rodlets (cf. p. 623), the lamellar planes corresponding to the 101 planes, they calculated how the spacial distribution of the various planes of the crystallites would vary as a function of elongation taking the observed volume changes into consideration. The calculation showed that the orientation of the 101 planes will run ahead of the orientation of the 10$\bar{1}$ and 002 planes. Evaluation of X-ray photographs revealed that this actually occurred and, moreover, they found that the overall rate of orientation as a function of the equivalent elongation v_a was approximately in conformity with the calculations.

Later, however, HERMANS and coworkers, using an improved and more accurate technique for the quantitative evaluation of X-ray photographs[2], could confirm the first result (the more rapid orientation of the 101 planes) but found that the overall orientation proceeds about twice as rapidly as predicted by theory. Moreover, the rate of orientation proved to be dependent on the composition of the cellulose solution from which the gels were prepared and, hence, to be associated with the special structure of the gel in question.

Both findings go to show that even in the process of elongation the crystallites are not simply dragged along with the matrix in which they are embedded but that quite another mechanism must prevail.

A further argument for rejecting the theory of affine deformation is, that, although it takes into account the observed changes in volume, it is incapable of explaining, let alone predicting, their occurrence.

According to the picture of gel structure prevailing at the present moment, the crystalline component of the gel does not consist of individual particles embedded in a matrix, but represents the permanent junction points in a coherent network structure formed by both the crystalline and the non crystalline component. This leads to the conclusion that the problem of the deformation mechanism will not be solved unless attention is paid to the properties of the network as a whole.

Recent work seems to indicate that this approach to the problem actually leads to more satisfactory results. The two principal points to be considered are
a) whether the observed course of the orientation in function of elongation is in conformity with the deformation of network structures,
b) whether network structures are capable of accounting for the observed phenomenon of anisotropic shrinkage without appreciable change in orientation.

To simplify the rather complicated problem, attention is shifted from the crystalline component to the non-crystalline one, the former being merely regarded as playing the part of juntion points in the network interconnected by the molecular fringes. The spatial extension of the crystallites and their movement on deformation are left out of consideration. This seems permitted since the crystalline component represents the minor portion of the substance in regenerated cellulose gels.

[1] *Vide* ref. 2 on previous page.
[2] P. H. HERMANS, J. J. HERMANS, D. VERMAAS, and A. WEIDINGER, *J. Polymer Sci.*, 1 (1946), 393; 2 (1947) 632; 3 (1948) 1.

Since attention is now focussed on the amorphous portion of the gel, X-ray analysis can no longer serve to obtain a quite correct measure of orientation and the latter should be derived from measurements of optical anisotropy. We have already seen on p. 638 that in the process of orientation by stretching, the optical orientation factor lags behind that derived from X-ray analysis. On page 593 it has been explained that the optical orientation factor is a measure of the average orientation of both the crystalline and amorphous components. However, since the latter forms the major portion of the gel, it will mainly determine the optical orientation factor. In some favourable cases the contribution of the crystalline component can, moreover, be computed and corrected for.

The general theory of the deformation of molecular networks has been referred to on page 634 where equations (57) (58) and (60) give the theoretical relation between optical anisotropy and elongation. In this theory it is assumed that all the distances between the junction points, that is the endpoints of the molecular chains forming the network, change on deformation according to the principle of affine deformation. Now, if we replace each molecular chain in the network by a vector *r* between its two endpoints, as dia-

Fig. 100. Molecular chains substituted by a vector *r* between their endpoints.

grammatically indicated in Fig. 100, the network can be described as a system of vectors which, upon deformation, change their spatial orientation according to the rules of affine deformation just in the same way as the rigid rodlets shown in Fig. 99 do, apart from the important difference that they also can become longer or shorter according to the very same rules.

From *Kuhn*'s calculations (p. 634) it follows that the anisotropy of the molecular chain in the direction of the vector increases or decreases with the length of the vector. Its contribution to the anisotropy of the filament (whose axis is supposed to lie in the OX direction) is hence entirely given by the length *r* and by the angle of orientation *a* of the vector.

Since it is assumed that any volume element of the gel is deformed in the same proportion as the macroscopic body, we can now picture what will happen to a given vector either on deformation or on shrinkage of the gel, and so obtain the analogy to the case of rigid rodlets represented in Fig. 99. This is done in Fig. 101, where *a* represents the vector in a volume element of an isotropic swollen gel. If the gel is dried isotropically (*c*) the vector will not change its orientation, but it becomes shorter. However, this does not give rise to a change in anisotropy since random distribution of all the vectors is maintained. Picture *b* represents the situation after a certain elongation of the body *a*. Upon elongation, the vector becomes longer and also at a smaller inclination relative to the direction of elongation. Both factors contribute to increased optical anisotropy. Fig. 101*d* shows what will happen to the vector on drying after deformation of the swollen gel. (For easier comparison the situation *b* is repeated in dotted lines). Here it is important to note that the network picture can qualitatively account for the constancy of anisotropy of the gel under the conditions of anisotropic shrinkage actually observed (which was impossible in

the case of Fig. 99). For, the vector becomes shorter, but also its slope smaller.

If the stretched anisotropic gel *b* happened to shrink isotropically, the vector would maintain its orientation a_2 and become only shorter. This would give rise to a decrease of anisotropy on drying. The same would apply if the drying entailed a greater relative lengthwise contraction than lateral contraction. The angle of orientation of the vector would then become larger and its length shorter. Both factors would contribute in reducing the anisotropy.

Only the actually observed case of a lateral contraction surpassing the lengthwise contraction can lead to constancy of anisotropy, since then $a_3 < a_2$ while the length of the vector decreases. These two factors have an opposite effect on the anisotropy of the system, so that they may compensate each other under suitable conditions. It could be shown that the anisotropy of swelling observed at various degrees of orientation can be semi-quantitatively accounted for on the basis of this picture.

The second of the two questions mentioned on page 642 can therefore — at least qualitatively — be answered in the affirmative and, moreover, the network theory leads to a reasonable explanation of the anisotropy of swelling.

Fig. 101.   Diagram of the changes in orientation and length of molecular vectors on deformation and shrinkage

We shall now briefly consider the first of the two questions. HERMANS and VERMAAS[1] have shown how the theory of KUHN can be easily modified for the case that volume changes also occur on stretching. They further collected extensive experimental material on the course of birefringence of cellulose filaments when stretched at various degrees of swelling. These data could serve for comparison with the theory.

From the slope of the optical anisotropy versus elongation curves the average number N of statistical chain elements per chain in the network can be estimated and is then found to be very small. Depending upon the various special cases examined, it lies in the order of $N = 1$ or 2, in conformity with the relatively small maximum extensibility of cellulose filaments of the order of 2 to 3, which, according to equation (56) on page 633, would also lead to an estimated value of N in the order of only one. This would indicate that cellulose gels behave as network structures consisting of very short chains or, in other words, that the junction points in the gel are very close together, even when taking into consideration that the cellulose molecule has a rather stiff chain and that, consequently, its statistical chain element is considerable longer than for instance that in rubber.

Though the authors mentioned found that several features in their experimental material on the behaviour of cellulose gels (which would led us to far afield to discuss here) seemed to conform with the general picture of molecular networks, a convincing quantitative interpretation on the basis of KUHN's theory, which is only valid for large values of N, remains uncertain if the chains are short.

---

[1] P. H. HERMANS and D. VERMAAS, *Trans. Faraday Soc.*, 42 B (1946) 155.

A further step in the development is the application of the theory of short chain network structures by J. J. HERMANS, which was already briefly mentioned on page 635, though also here the consideration of the case for N as small as 1 or 2 does not remain without certain difficulties. We shall not go into the details of the subject here, but refer to the literature[1], and confine ourselves to stating that the experimental data on the course of orientation can be reasonably well accounted for by the short chain theory.

It also leads to certain quantitative predictions on the stress-strain relations of filaments taken at various degrees of swelling which appeared to be particularly well fulfilled[1]. It is further capable of explaining in first approximation the course of the volume changes observed on stretching swollen filaments at various degrees of swelling.

Finally it should be emphasized that this theory, though in terms of somewhat modified concepts, is, in principle, also consistent with the general picture of deformation and swelling quite roughly characterized by the diagram given in Fig. 97 page 639. If freshly prepared swollen isotropic filaments are allowed to shrink to a lower degree of swelling before being stretched, their orientation versus elongation curves actually behave as though the average number of chain elements N were larger than that in the higher swollen state. This points to a higher degree of kinkyness of the chains which — as J. J. HERMANS showed — has the same effect as an enhanced value of N.

Fig. 102. Empirical relation between orientation derived from birefringence (f_o) and from X-ray photographs $f(x)$. The different signs refer to a variety of cellulose filaments prepared in a different way and covering swelling degrees between 16 and 2.2.

There remains hardly any doubt that the concept of a molecular network structure in cellulose gels is an essentially correct one.

None of the molecular network theories thus far discussed, however, is capable of predicting the course of the orientation of the crystallites which form the junction points of the network. It is only known from experiment (see p. 624) that the orientation of the crystallites runs ahead of that of the non crystalline gel component. From a large experimental material covering cellulose gels of various swelling degree and of different preparation, it was found that there seems to be a uniform relation between the optical orientation factor and that derived from X-rays holding for all cases. This relation, which may be of value for future theoretical considerations, is shown in Fig. 102.

---

[1] J. J. HERMANS, *Trans. Faraday Soc.*, 42 B (1946) 160, and ref. 2 p. 640.

It may be of interest finally, to show in a diagram the course of the orientation factor as a function of elongation according to the various theories briefly discussed in the foregoing (Fig. 103). The experimental values found for cellulose gels lie close to the $N = 1$ curve though the results vary somewhat with the degree of swelling and the previous history of the gel.

Our considerations may have served to illustrate that the general trend of research in this field is now more and more in the direction of particularly focussing attention on the rôle played by the *amorphous* portion of the structure. In this respect cellulose research will now probably follow a path more parallel to that of rubber research than hitherto has been the case. This parallelism between the extremes: rubber and cellulose will also arise from the problems referred to in next section.

No adequate theory has thus far been discovered capable of accounting for the orientation of the crystalline component. Doubtless it will be a very difficult, if not an insoluble, mechanical problem to connect the theory of network structure, which entirely abstracts itself from the nature, the possible spatial extension and the shape of the junction points, with the changes in the orientation of the latter.

Fig. 103. Theoretical orientation curves; I rigid rodlets following affine deformation of the matrix in which they are embedded (KRATKY-theory). Other curves: theory of molecular network with short chains. At each curve the number of statistical chainelements N per chain to which the curve refers is indicated.

9d *Comparative consideration of the mechanical properties of rubberlike substances and those of cellulose gels in connexion with the mechanism of deformation*

α. Generalities. In the foregoing sections, dealing with rubber and cellulose, the idea has been put forward that in both cases the mechanism of deformation essentially consists of the deformation of a network-structure. The concept of a network structure implies that only a micro-flow process will occur and macro-flow is excluded. If the mechanical behaviour of these substances should be represented by a mechanical model, consisting of springs and dashpots of the kind discussed in Chapter I of Vol. I, it should, hence, be a model in which any dashpot has at least one spring parallel to it.

The factors inhibiting macro-flow in vulcanised rubber are the chemical cross links due to vulcanisation (and probably also molecular entanglements), considerably assisted by the phenomenon of crystallisation. In unvulcanised raw rubber macro-flow may occur to a certain extend. In cellulose crystalline junction points of a high degree of stability are responsible for the exclusion of macro-flow. In rubber the chains are very flexible and consist of a large number of statistical chain sections; in cellulose the chains are stiffer and consist of a small number of chain elements. In the former case the intermolecular forces are weak, in the latter case they are strong.

Permanent elongation imposed on an originally isotropic macro-molecular system by some mechanical treatment, is by no means an indication of macro-flow if the system is anisotropic in the final condition. Though the systems under consideration will generally tend to reassume the isotropic condition and recover their original dimensions, the recovery may be *blocked*. This *blocking* is due to molecular cohesion. It may often be more or less completely released by such means which lessens intermolecular cohesion. We then observe the phenomena designated as thermo-retraction, swelling-retraction and the like. The rubbery condition is characterized by a minimum of blocking; cellulose, on the contrary, is an example of a substance where blocking-phenomena play a decisive part in deformations. In the mechanical spring and dashpot models blocking should be represented by a special device, inhibiting extended springs from retraction, either entirely or over a part of their length.

The spring action may be associated with the molecular entropy springs, which, however, may also more or less assume the character of energy-springs if extension of main valency bonds or deformations of valency angles is involved in molecular deformation. The dashpots are associated with the internal frictions, say, the viscosity, of the body[1]. Molecular springs will the more readily assume the character of energy springs, the greater is the internal friction. The latter, again, will be the more considerable the greater is the intermolecular cohesion.

With these general considerations as a tool, a great deal of the phenomena concerned with the mechanical properties of macro-molecular systems can be qualitatively interpreted and understood.

β. The stress-strain relations of rubber-like substances and that of isotropic cellulose.

It has already more than once been said that, owing to its high intermolecular cohesion, dry isotropic cellulose is a glass-like inextensible substance, just as rubber
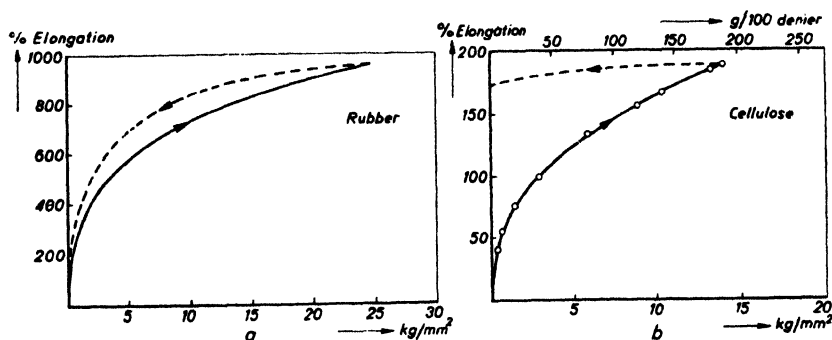
Fig. 104. Stress strain diagrams of (a) swollen isotropic cellulose filaments and (b) of moderately vulcanized rubber at loading and unloading (stress on actual cross-section)[2].

---

[1] This viscosity might, of course, in its turn again be represented by a series of little spring dashpot systems of the type shown in Volume I, Chapter I, all having a very short relaxation time.

[2] In swollen cellulose the weight of dry substance per unit length of the filament was taken and the cross section calculated from that figure.

is when cooled in liquid air. That isotropic cellulose is an extensible substance under common conditions is due to its high affinity towards water. Through the absorption of water in the amorphous regions intermolecular cohesion in the latter is weakened, since the cohesive forces are, so to say, partially screened by the absorbed water molecules. In highly swollen cellulose gels, where the chains of the network lie at relatively large distances from each other, the freedom of their motion is maximum.
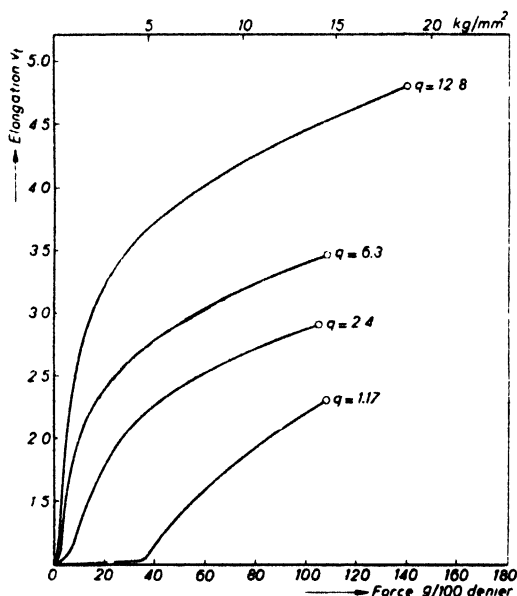
Fig. 105. Stress-strain diagrams of isotropic cellulose filaments at various degrees of swelling q. (stress on actual cross section).

Fig. 104a and b show that, apart from the absolute value of extensibility, the stress strain curves of highly swollen isotropic cellulose filaments are very similar in shape to those of isotropic rubber[1], in as far as the curve taken on loading is concerned. The difference may be ascribed to the smaller number of chain sections N in the former.

Upon unloading, the rubber shows complete recovery. (The magnitude of the hysteresis between loading and unloading depends on temperature but is not very large). The cellulose filaments show a very imperfect recovery. The latter is *blocked* in that new secondary junction points have been formed between the molecules in their new orientated positions. Rubber would show the same behaviour if sufficiently cooled in the extended state.

It is now of interest to see the influence of the degree of swelling of the isotropic cellulose on the stress-strain relation. This is demonstrated in Fig. 105[2]. According as the swelling of the objects is lower, a new phenomenon enters the picture to a more and more marked extent: the appearance of a *yield value*. Elongation does not begin before a certain stress has been surpassed. This may be interpreted as another kind of blocking. The overall deformation is essentially that of the primary network structure as determined by the primary junction points of the gel. On de-swelling and particularly on drying to the air-dry condition, the interchain distances become smaller and smaller. Consequently, new secondary junction points are locally formed between the chains. These secondary junction points must be disrupted before the deformation of the primary network as such can take place. An initial stress is required to overcome this blocking, which we might term *drying-blocking* in contrast to the before mentioned blocking, which we then might term *orientation-blocking*.

---

    [1] P. H. HERMANS, *Kolloid-Z.*, 86 (1938) 107; *Proc. Roy. Acad. Sci. Amsterdam*, 43 (1940) 1032; *J. Phys. Chem.*, 45 (1940) 827.
    [2] For reasons elsewhere explained (*Proc. Roy. Acad. Sci. Amsterdam*, 42 (1939), 798, *Kolloid-Z.*, 89 (1939), 344) the ,,rational" measure of extension v_t is used in this figure.

It need hardly be said that a similar kind of blocking will occur in rubber when lowering the temperatures. Fig. 106 shows the stress-strain curves of a rubber sample taken at decreasing temperatures according to FUJIWARA and TANAKA[1]. The analogy with the cellulose curves at decreasing degree of swelling is evident. It is also of interest to compare the stress-strain diagrams of balata and gutta percha, polymeric hydrocarbons whose temperature of crystallisation lies considerably higher than that of ordinary rubber.

Fig. 107 refers to balata. This substance is distinctly crystalline at ordinary temperature

Fig. 106. Stress strain curves of rubber on lowering the temperature.

Fig. 107. Stress strain diagrams of balata, I at ordinary temperature, II at elevated temperature on loading and on unloading.

and it shows a marked yield value: Before the "molecular springs" can come into action, they must first be torn loose from each other.

The junctions which have to be overcome initially are quite comparable to those occurring in air-dry cellulose. Beyond the yield point the balata curve takes the usual course, typical for molecular network structures. At high degrees of extension, just as in rubber, crystallisation sets in again. (Orientation-blocking). That is why recovery is very imperfect upon unloading. If the temperature is now sufficiently raised, almost complete recovery occurs. At higher temperatures, gutta percha and balata behave like rubber.

Hence, two kinds of blocking make their appearance together with increased molecular cohesion. The first one, orientation blocking, is always connected with a certain degree of anisotropy of the body concerned. It leaves a "memory" for the deformation which has previously taken place. Its release is connected with a recovery. The other one (drying- or crystallisation-blocking) is not necessarily connected with anisotropy. It may ocur

[1] T. FUJIWARA and T. TANAKA, Rubber Chem. and Techn., 7 (1934), 610; cf. also M. LEBLANC and M. KRÖGER, Kolloid-Z., 37 (1925), 205.

in perfectly isotropic bodies as well. It will be clear that, in orientated bodies, both kinds of blocking will play a rôle.

γ. Stress-strain relations and orientation. The stress-strain relations of cellulose fibres mentioned in the literature always refer to orientated objects. It has little, if any, sense to compare them with those of other, isotropic, substances, as e. g. rubber, and then discuss the difference as being characteristic for both classes of substances[1]. This has been frequently overlooked. We certainly follow a more rational course in studying how the stress-strain relation of cellulose changes in dependence of its initial orientation. Fig. 108 shows the stress-strain curves of a series of cellulose filaments which were prepared by taking freshly prepared cellulose xanthate filaments, stretching the latter to increasing extents and then decomposing and drying the
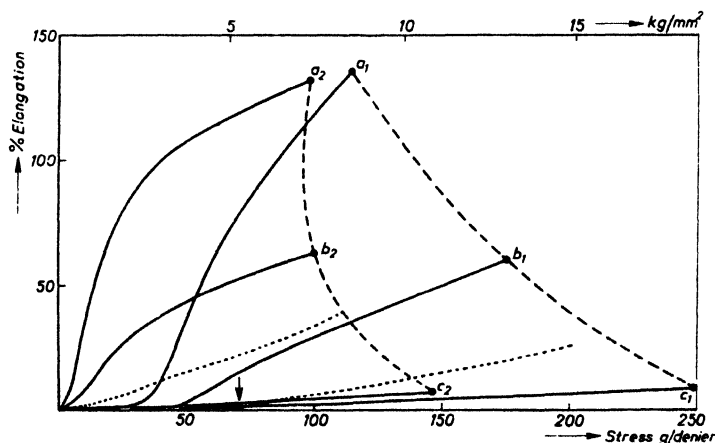
Fig. 108. Stress-strain diagrams of model filaments in the dry (subscript 1) and in the wet (subscript 2) condition; a isotropic filaments; b and c filaments previously stretched 50 and 100 per cent. in a highly swollen condition (dotted curve: ordinary viscose rayon).

orientated filaments obtained. The stress-strain curves of these filaments, taken in the air-dry condition as well as after being allowed to swell in water, are both shown in Fig. 108. First looking to the dry filaments it is seen that the curves gradually change their shape. The yield value increases with the orientation. This will be clear, since the blocking which has to be overcome initially is now not merely the blocking due to drying. The orientation-blocking to which the orientated fibre was subjected at the previous process of orientation and which prevented it from complete recovery will also contribute a component to the total amount of blocking. If the deformation has taken place over a certain range, and the fibre is then released, it enters a new state of blocking and remains there. The object is, so to speak, transferred, through a release, from the first state of blocking to another one. In other words, existing secondary junction points are released and new ones are formed in another

---

[1] Cf. P. H. HERMANS, *Kolloid-Z.*, 86 (1938), 107.

pattern. The mechanical behaviour of the majority of textile fibres on loading and unloading may be conveniently interpreted on this basis[1].

Subsequent swelling in water of a filament stretched in the air-dry condition partly removes the blocking and a considerable retraction is observed. In cellulose fibres, which were already moderately or highly orientated in the initial condition, like rayon filaments, this recovery is even almost a complete one. To bring about this swelling-retraction, a degree of swelling must be reached which is higher than that at which the orientation has been performed[2]. Hence, swelling in water does *not* give rise to a recovery of the orientation imposed on the fibres by their previous stretching in the highly swollen xanthate state.

The stress strain curves of the swollen filaments shown in Fig. 108 do not give rise to many additional remarks. The yield values are, for reasons obvious now, considerably lower than those of the air dry filaments.

The broken lines in Fig. 108 indicate how the tensile strength of the filaments changes as their orientation proceeds[3]. The dotted lines are the stress-strain curves of an arbitraryly selected viscose rayon filament which fits in well with the general scheme.

From a qualitative point of view, the mechanism of deformation of macromolecular substances seems to be rather well elucidated at present. In quantitative respect, however, numerous very difficult problems remain to be solved.

---

[1] We cannot go into further details here. The subject has been recently set forth by H. A. VREE-DENBERG, *J. Polymer Sci.*, 1 (1946), 329. The author owes to DR. VREEDENBERG several stimulations in the development of his own ideas.

[2] P. H. HERMANS, *Cellulosechemie*, 19 (1941) 117.

[3] For a more detailed discussion of the mechanical properties of cellulose, cf. ref. 2 on p. 640.

XIII. SOLID MACROMOLECULAR SYSTEMS
WITH ONE (CHEMICAL) COMPONENT

by R. HOUWINK

Rubber Foundation, Delft

---

§ 1. INTRODUCTION

In this chapter solid systems containing only one kind of macromolecule, from a *chemical* point of view, will be treated. Thus *isogels*, already mentioned on p. 44, containing usually a more or less continuous series of molecules of different size, will be met with here. Substances such as resins and synthetic rubbers may be mentioned as typical examples.

There is a group of materials closely connected with these systems, namely, those *mixtures of isogels* containing two or more chemical components, each consisting of a more or less continuous range of molecules of different size (glass, asphalt bitumen). Strictly speaking, they belong to the group of *heterogels*, treated on page 42 which are characterised by the presence of at least two different chemical components. Mixtures of isogels, however, differ from heterogels proper, in so far that the latter possess at least one component having molecules of a certain size only (water in a gelatine gel). Mixtures of isogels will also be considered in this chapter, because their properties are much similar to those of pure isogels.

We will see on p. 654, that there is a gradual transition of physical properties on passing from isosols, via isogels, to hard solid systems, substances which appear to be like well cooled glass or hardened resins, where freely rotating ("liquid") molecules are no longer present.

It will appear in this chapter that the transition, sol → gel → hard solid state, can be accomplished by cooling, and also by polymerisation or by interlinking, this being the reason for the special importance of these phenomena in polymer chemistry.

§ 2. OBSERVATIONS ON COOLING A MACROMOLECULAR LIQUID
a. Transition points

Let us first consider the cooling of a non-crystallising liquid, such as molten asphalt. In such a liquid the particles may have a high degree of association, which increases still further on cooling. As soon as a certain coherence has been established all through the material, the substance becomes a solid with a yield value[1], but it still contains freely rotating particles, hence it is called a gel. The transition temperature

---

[1] Instead of a yield value, a viscosity coefficient greater than 10^5 poises, measured at a shearing stress of 10^3 dynes/cm², is considered as a criterion for the solid state on practical grounds.

between the sol and the gel will be denoted as the *fluidity temperature* T_f (see Fig. 1).

On further cooling, a hard rigid state is obtained at a certain temperature, which is called the *brittle temperature* T_b. It may be explained by assuming that the remaining freely rotating molecules also become fixed. From this supposition, and by analogy with the freezing point of crystals, it is a first-order transition point.

When cooling such brittle macromolecular substances further, another transition point can be observed, the *second-order transition point* T_m. It may be supposed that here the oscillations of certain molecules around their equilibrium positions become stopped, just as is observed with low-molecular crystallising substances below their melting points[1].

Summarising therefore, a contrast may be made between polymers and crystallising low-molecular substances like NaCl, in that all liquid molecules in the latter become attached suddenly at a certain temperature to form the crystalline lattice, whereas with polymers the corresponding change occurs between T_f and T_b, between which a gel state exists.

The above idealised picture for polymers is much more complicated[2] in practice. For linear polymers, the macro- and micro-Brownian motions with all their consequences must be considered. Taking into account the various factors involved, differences in rotational possibilities along the chain axis, differences in mutual attraction and in bulkiness of side groups, and distribution of molecular length, it is felt impossible to give one general picture covering all the phenomena to be observed. There is evidence that the completely free macro-Brownian motion of the molecules becomes hampered at T_f, and that between T_f and T_b, at least very large segments have a certain freedom of motion, and a large degree of rotation which is more or less stopped at T_b. Between T_b and T_m some further restrictions of motion will take place, and it is assumed that below T_m the segments can only oscillate about equilibrium positions at right angles to their length. T_b has been designated an "external melting point", and T_m as an "internal melting point".

The existing confusion is increased by the fact that various physical properties can be criteria for measuring the transition points: the heat capacity, the volume change, the thermal conductivity, the dielectric loss factor, the modulus of elasticity and the elastic extensibility are examples. It is, however, clear that not all these properties will show a sudden change at exactly the same temperature, because each of them will be connected in a somewhat different way with the molecular structure. For this reason it is not easy to quote reliable data for the transition points, a difficulty which is enlarged by the fact that such transition points are never very sharp, in contrast to the melting points of crystals. Considering only the micro-Brownian movements for instance, most polymers are characterised by a spectrum of motions. Each type will have its own relaxation time λ, which according to KUHN[3] is connected with its modulus of elasticity E and its viscocity η in the following way:

$$\eta = 0.385 \, E \, \lambda \tag{1}$$

[1] Thus, it has been proved (J. M. BIJVOET and J. A. A. KETELAAR, *J. Am. Chem. Soc.*, 54 (1932) 625) that in sodium nitrate with a melting point of 310° C, practically all nitrate groups still rotate at 275° C.

[2] For a general survey see: R. F. BOYER and R. S. SPENCER, *Advances in Colloid Science*, Vol. 2, New York, 1946.

[3] W. KUHN, *Z. physik. Chem.*, B 42 (1939) 1.

On cooling, the different groups will gradually come to a standstill; those with the greatest relaxation time first.

In spite of these objections[1], the transition points are very helpful tools in clearing up the picture of polymers, and in our opinion they will become even more important still, with a greater perfection of methods for their measurement. It will appear in the next section that a need for introducing other types of transition point (T_r, T_K) already exists.

b. The transformation interval

The temperature range between the limits (T_f and T_b) of the gel state is denoted as the *transformation interval*, this usually being broader for more highly polymerised substances. It is of the greatest technical importance, because in this interval are encountered the conditions necessary for the existence of rubbers and of soft plastics.
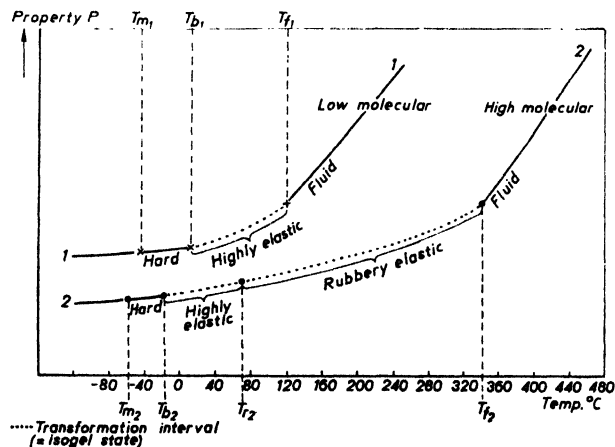
Fig. 1. *Transformation interval for low and for high polymers* (T_f = temperature of fluidity; T_b = brittle point)

If some representative physical property (e.g. density or elastic extensibility), indicated in a general way by P, is plotted against temperature, the situation shown in Fig. 1 is obtained[2]. Following the curve for the comparatively low-molecular material I and starting from such a high temperature that the material is in the sol state, cooling leads first to a temperature T_{f1}, where the fluid state ceases to exist and the solid (gel) state appears, the material becoming highly elastic at this point. Further cooling leads to T_{b1}, the brittle point where the highly elastic (gel) state ceases to exist, the material becoming rigid and hard. Expressing this in rheological terms[3], it may be said that the material is viscous above T_{f1}, highly elastic between T_{f1} and T_{b1}, and purely elastic below T_{b1}. Still further cooling leads to T_{m1}. No sharp distinction between the type of brittleness at T_{b1} and at T_{m1} has been made as yet. Presumably T_{m1} will correspond to some kind of super-brittleness.

For the high-molecular material similar regions can be distinguished, the transformation interval being much broader however. Experience has shown that this

---

[1] Further on BOYER and SPENCER have clearly shown the speed of testing to be essential in measuring the brittle point.

[2] Cf. E. JENCKEL und K. UBERREITER, Z. phys. Chem., A 182 (1938) 361. The original considerations of this type have been given by G. TAMMAN, *Der Glaszustand*, Leipzig, 1933.

[3] For the definitions see KRUYT in Volume I of this book.

interval can be subdivided into two parts: one, where the material is of a rubbery[1] elasticity (elastic extensibility $>$ 100%) the other, where it is only highly[1] elastic (elastic extensibility 1—100%). The transition of these two parts into each other is indicated by $T_{r,s}$, to be called the freezing temperature of the rubbery state[2].

There are reasons for assuming that at T_b the viscosity for any material is of the order of 10^{13} poises[3]. As already remarked, the higher the molecular weight, the broader usually is the region between T_f and T_b, a fact which may be illustrated by the following data[3].

TABLE 1

TRANSFORMATION INTERVAL FOR SOME SUBSTANCES (IN °C)

		T_f	T_b	$T_f - T_b$
Colophony	low-molecular	53	30	23
Shellac		57	31	26
Polystyrene	high-molecular	138	28	110
Polyindene		43	3	40

T_r, the temperature at which the highly elastic state ceases to exist, has been the subject of many systematic studies[4]. Of course, it is closely related to the chemical structure of the polymers involved, although the situation is often not easily explained, see p. 660.

It will appear later, that many polymers show a strong *crystallisation* tendency on cooling, which can be increased on stretching the polymer. The quality of some rubbers is often decided by this tendency (see p. 666).

c. Examples

First, we will consider the phenomena connected with the transformation interval of some materials with three-dimensional molecules (glass, asphalt), turning later to a chain polymer (sulphur), where the phenomena are much more complicated.

1°. Glass

Perhaps it may seem unusual that the isosol—isogel theory should be applied to glass. Considering however the modern theories about the constitution of glass[5], little doubt is left that in the transformation interval a network is present, consisting of ions cohering by primary forces, while at the same time parts of this net are freely rotating, according to Fig. 2.

Strictly speaking, most technical glasses are not pure isogels, however. Most of them contain various kinds of ions ($Na^·$, $K^·$, $Ca^{··}$, SO''_4, SiO''_3) with different

---

[1] For the definitions see KRUYT in Volume I of this book.
[2] Most authors do not discriminate between T_r and T_g, but consider T_g as the freezing temperature of the rubbery state. This is not exact however, because there is a transition between the rubbery elastic and the rigid state, namely the highly elastic state (elastic extensibility between 1 and 100%).
[3] G. TAMMANN, *Der Glaszustand*, Leipzig, 1933.
[4] Although it must be said that T_b has usually been considered instead of T_r.
[5] W. H. ZACHARIASEN, *J. Am. Chem. Soc.*, 54 (1932) 3841; *J. Chem. Phys.*, 3 (1935) 162; N. W. TAYLOR, E. P. McNAMARA and J. SHERMAN, *J. Soc. Glass Techn.*, 21 (1937) 61; G. HÄGG, *J. Chem. Phys.*, 3 (1935) 42.

bonding energies, and therefore, are not loosened at the same temperature. One could say that Na_2SiO_3 for instance, will form an isogel and the same can be said for $CaSiO_3$. This leads to the situation indicated above as a *mixture of isogels*. The network is not regular and crystalline as that pictured in Fig. 2a but it shows an irregular arrangement, leading to the amorphous structure of Fig. 2b.

As all ions in Fig. 2b do not have an equal position, the energy necessary to loosen them will be different, the melting thus taking place over a certain temperature

    a. Crystalline   *b.* Amorphous

O   = Oxygen atom
●   = "A" atom
- - - = "Liquid"
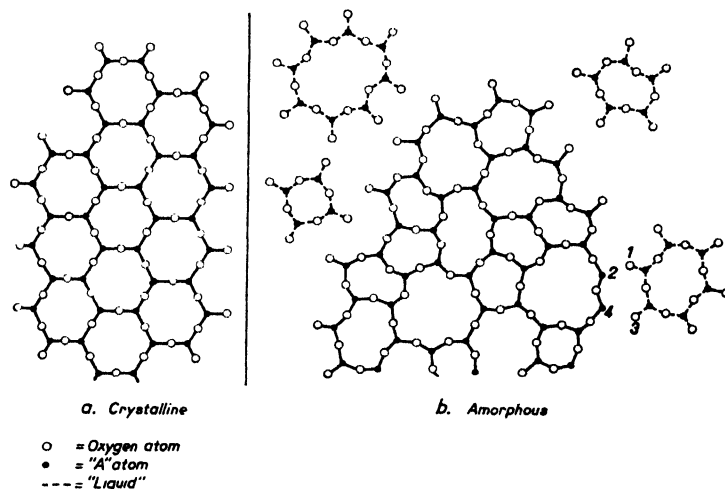
Fig. 2. Crystalline and amorphous arrangements of A_2O_3 in glass (partly after ZACHARIASEN). For the crystalline pattern the angle A-O-A is always the same, for the amorphous arrangement this angle changes from O-atom to O-atom.

range, the transformation interval. LILLIE[1] even calculates with the aid of formula (16), p. 182 that on cooling glass from 1100 °C to 500 °C, the weight of the aggregates increases from 15 000 to 112 000, the latter aggregate possessing a diameter of about 50 Å and being therefore of colloidal dimensions. In accordance with this gel theory, TAYLOR[2] has found that glass shows highly elastic properties in the transformation interval. When lowering the temperature, these were accompanied by elastic after-effects (see p. 664); this is in agreement with the expectation that the increased viscosity retards recovery. The existence of a certain coherence between the molecules in the transformation interval is also confirmed by the fact that quasi-viscous flow (see Volume I) is observed. Fig. 3 shows the change of the viscosity coefficient for selenium glass[3] on applying increasing stresses, caused by a gradual loosening of the aggregates from each other.

The fact that on cooling, glass changes from a gel into a rigid solid state needs

[1] H. R. LILLIE, *J. Am. Ceram. Soc.*, 16 (1933) 619.
[2] N. H. TAYLOR and co-workers, *J. Soc. Glass Techn.*, 21 (1937) 61.
[3] E. JENCKEL, *Kolloid-Z.*, 84 (1938) Part 3.

no further elucidation. At room temperature it shows pure elasticity with an extensibility of less than 1%, and thus it behaves just like crystalline substances, such as metals.

Its modulus of elasticity is also of the same magnitude as that of crystals. The tensile strength of roughly 500 kg/cm² is about 200 times lower than might be expected on theoretical grounds. This discrepancy is explained by SMEKAL [1], who assumes small chinks (German: Lockerstellen, meaning "weak spots") to be present throughout the material. In our opinion [2] the formation of such chinks is easily understood on a basis of the isogel theory developed above. Considering Fig. 2b, it is to be expected that on cooling the isogel, the "liquid" aggregates will not always have the opportunity of setting

Fig. 3. Quasi-viscosity of selenium glass in the transformation interval (η_{100} means: η at a shearing stress of 100 kg/cm²).

themselves in the position of lowest potential energy, being too voluminous to rotate quickly enough in the highly viscous medium. So it might be that the oxygen atom 1 just manages to fix itself to the A-atom 2, while in the meantime the viscosity has increased so much as a consequence of the cooling process that the oxygen atom 3 is prevented from becoming fixed to the A-atom 4, thus leaving some free space, a small chink. A similar theory will be advanced on p. 673 for the hardening of resins, where the same discrepancy between theoretical and experimental strength is encountered.

2°. *Asphalt bitumen*

Another group of materials forming mixed isogels in the transformation interval and becoming hard solid substances on cooling further, are the asphalt bitumens. These contain [3] two groups of constituents of special importance for our considerations, namely, the higher molecular *asphaltenes* and the lower molecular *maltenes*. The asphaltenes are considered as high-molecular hydrocarbons of a predominantly aromatic or hydro-aromatic character. The maltenes may not only be aromatic but aliphatic as well; the more aromatic they are, the better will they peptise the asphaltenes.

It has been found that the type of flow depends chiefly on the properties and quantity of the maltenes. If these are aromatic, pure viscous flow is observed, indicating the peptisation of the asphaltenes to be to such an extent (Fig. 4a) that these do not hinder each others motions. If however the maltenes are aliphatic or if they

[1] A. SMEKAL, *International Conference on Physics*, London, 1935, Part II; *Ergeb. exakt. Naturw.*, Berlin, 1936, p. 173.

[2] R. HOUWINK, *Elasticity, Plasticity and Structure of Matter*, Cambridge 1937.

[3] The most far-reaching considerations about the structure of asphalt bitumen were given by J. PH. PFEIFFER and P. M. VAN DOORMAAL, *J. Inst. Petrol. Techn.*, 22 (1936) 414. Here also, the older work of MARCUSSON, NELLENSTEYN, MACK and others is reviewed.

are removed, a point is reached where the asphaltenes gradually adhere to each other, forming the skeleton shown in Fig. 4b (blown-asphalt[1]), and leading to quasi-viscous flow, high elasticity and even thixotropy.

⊕ Nuclei of the micelle (aromatic)
❍ Compound of high molecular weight (aromatic)
◯ Compound of a lower molecular weight (aromatic)
○ Compound of a partly aliphatic character
⌒ Compound of a still more aliphatic character
– Compound of a nearly complete aliphatic character
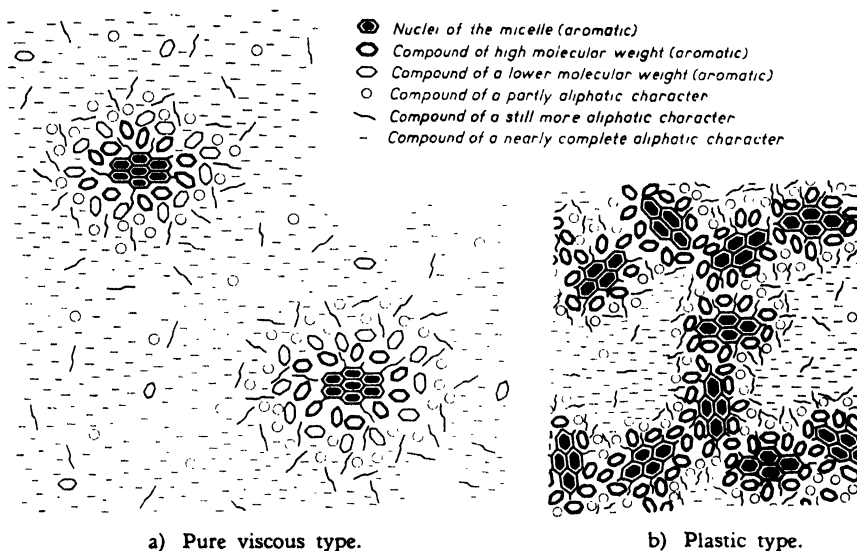
　a) Pure viscous type.　　　　　　　　　　　b) Plastic type.
Fig. 4.　Structure of asphalt bitumens[2]

It is obvious that on cooling bitumen, all Brownian motion of the maltenes also will finally be stopped, leading to a hard solid state. This is the brittleness of asphalt, well known in winter.

3°. *Sulphur*

It has already been stated (p. 38) that sulphur forms chain molecules on quick cooling after being kept at 250° C. If plastic sulphur, obtained on pouring the melt into cold water, is stretched, crystals are formed, built up from oriented polymers, composed of S_8-units[3]. Very pronounced, highly elastic properties and a considerable tensile strength can be observed at this stage, depending upon the temperature of preparation and upon the time elapsing after the cooling process.

This can be followed by considering the changes in the stress-strain diagram (Fig. 5). The material, tested immediately after being prepared at 250° C, is rather stiff; on applying a stress of 40 kg/cm² it shows an extension of only about 150%, whereas for the 350° C material this value is 500%. The stress-strain curve for the last-mentioned material shows a close resemblance to that for raw rubber, as appears by comparison with the dotted line in Fig. 5b. Upon storing, the highly elastic sulphur becomes stiffer.

[1] Here, maltenes have been converted into asphaltenes by blowing hot air through the material.
[2] J. PH. PFEIFFER, *De Ingenieur*, 21st July (1939) MK 41.
[3] K. H. MEYER, *Trans. Faraday Soc.*, 32 I (1936) 148.

It is not yet possible to give a reliable explanation of all these phenomena (when considering them in detail, still more complications are encountered [1]), but some indications can be borrowed from viscosity data which seem to lead to the key for unravelling these problems. The viscosity-temperature function of molten sulphur

a) Temperature of preparation, 250° C.

b) Temperature of preparation, 350° C.
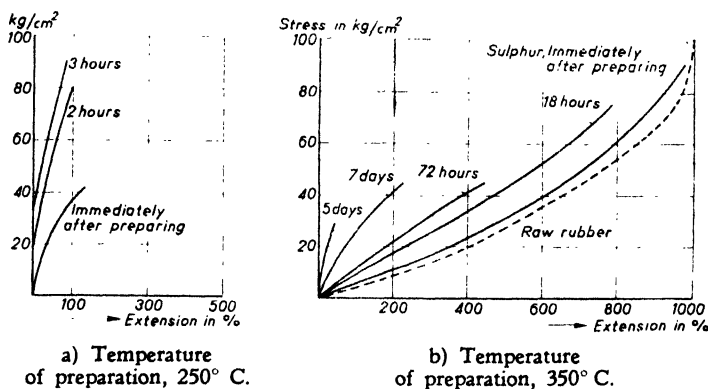
Fig. 5. Changes with time of the stress-strain diagram of sulphur prepared at different temperatures.

(Fig. 6) shows a sharp peak between 180 and 250° C. Referring to p. 169 for the influence of molecular length on viscosity, it is obvious that the formation of chain molecules must be considered responsible for this peak. We know from other examples (cf., cellulose p. 665) that a great chain length leads to a strong material and thus it can be explained that suitably prepared sulphur will have a considerable strength. The stiffening on storing may be due to crystallisation, bringing the chains into a state of lower potential energy, thus increasing their mutual coherence.

For the other inorganic substances with chain molecules mentioned on p. 38, rubbery elasticity can also be observed in the transformation interval.
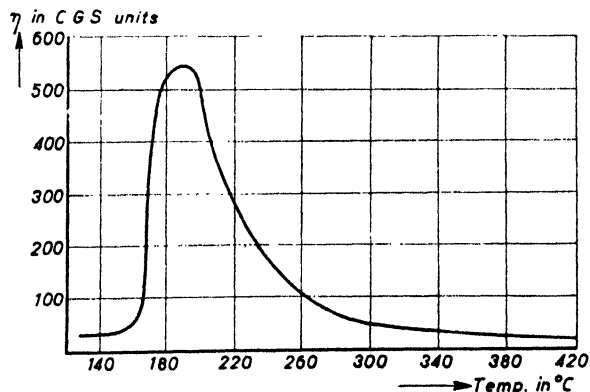
Fig. 6. Viscosity change of molten sulphur on heating [2]

---

[1] The tensile strength, for instance, changes in a very complicated way; cf. the original studies of K. SAKURADA and H. ERBRING, Kolloid-Z., 72 (1935) 129 and the survey by R. HOUWINK, Elasticity, Plasticity and Structure of Matter, Cambridge, 1937. For an exact explanation of all phenomena observed, much more experimental material is necessary. Thus, viscosity and tensile strength tests should be carried out with one and the same material. At the same time the extent to which the deformations are elastic or plastic should be studied.

[2] C. C. FARR and D. B. Mc. LEOD, Proc. Roy. Soc., London, A. 97 (1920) 80.

§ 3. LINEAR POLYMERS

a. The transformation interval

When a monomer or a fluid polymer is subject to linear polymerisation it will ultimately become a solid and, just as on cooling, it will also pass from the isosol, through the isogel, to the hard rigid state. The analogy can be expressed by means of the following equation for the kinetic energy:

$$\tfrac{1}{2}\, mv^2 = \tfrac{3}{2}\, kT \qquad\qquad (2)$$

On cooling, T decreases, and consequently the velocity v of the molecules diminishes, ultimately leading to a cessation of the Brownian motion. On polymerisation however the mass m increases, leading to the same result.

The structural changes for chain molecules can be represented by Fig. 7. One of the features of the isogels (Fig. 7b) of synthetic polymers is that, for the greater

<div align="center">

a) Sol   b) Isogel   c) One-phase system

</div>
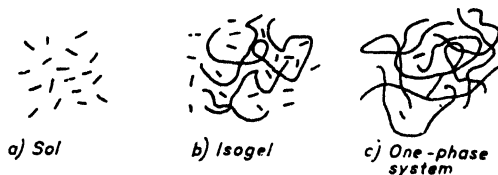
Fig. 7. Changes in colloidal structure on polymerisation of chain molecules.

part[1], only large molecules are present next to the small "liquid" molecules. The reason for this has been indicated on p. 43, the activated monomer molecules growing rapidly and immediately to their final length. In the final one-phase system some small molecules may be assumed to be still present, the chance of contact for the last traces of monomer being too small and tne viscosity too high for a complete finishing of the reaction.

For natural polymers the situation is probably different, the possibility being left open that the chain is built up directly from products like CO_2, formaldehyde, etc., without the formation of a monomer (p. 45). In this case, no isogel would be formed at all, especially when the polydispersion is less pronounced than for synthetic polymers, which for many natural products may be expected (p. 47).

Keeping in mind the restrictions, noted on p. 654 when quoting data for the transition points, some values will be mentioned in Table 2 which will be considered as "fair averages".

It does not appear easy to detect a simple rule, connecting the molecular structure with either T_m or T_b. Generally speaking the trend of T_m is similar to that of T_b, indicating that essentially, either of these temperatures is controlling the same physical property.

It might be expected that the trend is the same as that of the molecular cohesion of the side groups.

The last column indicates that for polymethyl acrylate and polyvinyl chloride, for instance, this is not true and therefore other factors must play a part.

---

[1] A certain polydispersion always exists, see p. 44.

TABLE 2

VALUES[1] FOR T_m AND T_b

	Second-order Transition temp. T_m in °C	Brittle point T_b in °C	Side group X of vinyl	Mol. Cohesion per group X kcal/mol
Polyisobutylene	—74	—50	$(CH_3)_2$	1.8
Polyisoprene (rubber). .	—73	—58	$<^{CH_3}_{CH \; - \; CH_2}$	1.5
Polyethylene	—	—68	none	1.0
Gutta Percha (pure) . .	—	—54	$-<^{CH_3}_{CH \; - \; CH_2}$	1.5
Neoprene GN	—	—40	$<^{Cl}_{CH \; - \; CH_2}$	—
GRS	—61	—30	—	—
Polyvinylidene chloride .	—17	—	Cl_2	—
Polymethyl acrylate . .	3	0 to 8	$COOCH_3$	5.6
Polyvinyl acetate . . .	28	—	$OCOCH_3$	3.9
Ethyl-cellulose	43	—	—	—
Cellulose nitrate	53	—	—	—
Polymethyl methacrylate	57—68	—	—	—
Cellulose acetate	69	—	—	—
Polyvinyl chloride . . .	75	81	Cl	3.4
Ebonite	80—85	—	—	—
Polyacrylic acid	80—95	—	COOH	8.9
Polystyrene	81	80	C_6H_5	—
Glyptal resin	83	—	—	—
Polyvinyl alcohol. . . .	85	—	OH	7.2
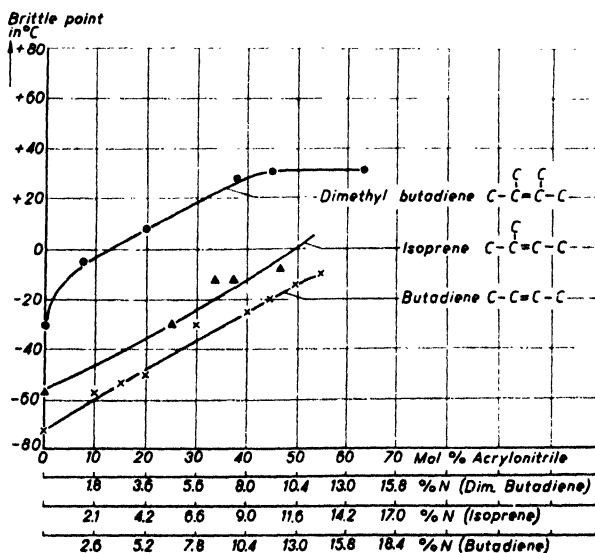
Fig. 7a. Shift of the brittle points of some dienes by introducing acrylonitrile.

[1] Values borrowed from the surveys of R. F. BOYER and R. S. SPENCER, *Advances in Colloid Science*, Vol. 2, New York, 1946 and from E. JENCKEL, *Kolloid-Z.*, 100 (1942) 163.

These are considered to be the bulkiness[1] of side groups (polystyrene[2]), the degree of symmetry (vinyl chloride compared with vinylidene chloride) and the energy barrier for rotation around the C-C bonds in the main chain[3].

When however the constitution of a molecule is changed in a very systematic way, a certain rule can be detected, as appears[4] from Fig. 7A. The first point to be observed is the upward trend of the brittle points of all dienes on introducing the strongly polar acrylonitrile. The second point is the influence of the methylene group. Butadiene without methyl groups is the most easily mobile molecule, polyisoprene with one methyl group is stiffer, dimethylbutadiene being the stiffest.

In practice this has worked out in so-far, that from dimethylbutadiene no valuable rubber can be made, whereas the two other dienes provide mobile snappy rubbers with a low brittle point.

b. Behaviour on deformation in the transformation interval

1°. High-Elasticity

Many chain polymers are of a rubbery elasticity in the transformation interval and the opinion is often expressed that the chain-form of the molecules is a necessary condition for this elasticity. This is not true however; on p. 672 it will be shown that rubbery elasticity can also be observed in materials with a three-dimensional structure. The conditions for rubbery elasticity would appear to be: a sufficient

a) Rubber   b) Gutta percha.
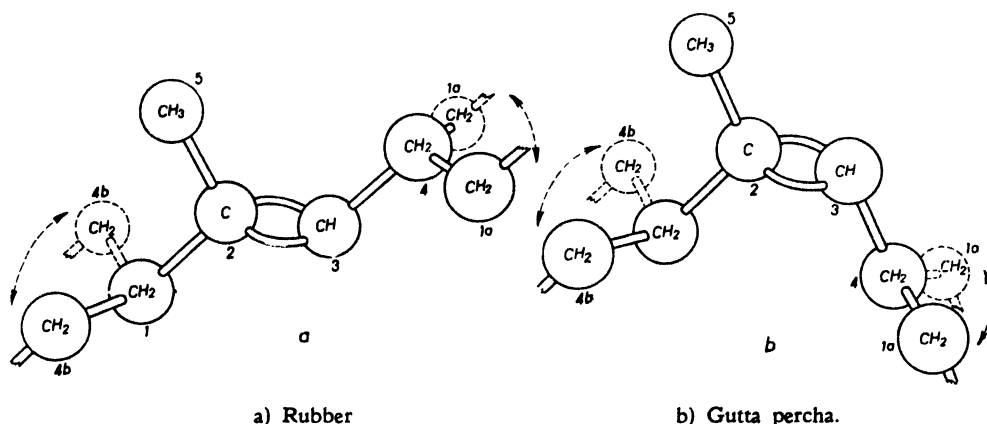
Fig. 8. Valence oscillations and rotations in rubber and gutta percha. (The diameters of the atoms have been designed too small by comparison, for reasons of distinctness) According to Bunn[5]

[1] H. T. NEHER, *Ind. Eng. Chem.*, 28 (1936) 267; cf. also *ibid.*, 28 (1936) 1160.

[2] Later J. H. DE BOER, *Trans. Faraday Scc.*, 32 (1936) 10 showed the steric conditions to be such that the benzene rings exert a strong attraction on each other.

[3] E. JENCKEL, *Kolloid-Z.*, 100 (1942) 163.

[4] A. M. BORDERS and R. D. JUVE, *Ind. Eng. Chem.*, 38 (1936) 1066.
The values for dimethylbutadiene are borrowed from unpublished work by G. SALOMON.

[5] This problem has been thoroughly investigated by K. H. MEYER, *Die hochmolekularen Verbindungen*, Leipzig, 1940, p. 116. His considerations have been largely extended by C. W. BUNN, *Proc. Roy. Soc. (London)*, A. 180 (1942) 40.

degree of polymerisation and a rather loose coherence between the molecules or the elements of the molecular nets; the structural units must also be able to rotate more or less freely.

For the purpose of illustration, the way in which structural factors of apparently minor importance may be decisive for the existence of high-elasticity [1] may be pointed out. Referring to p. 35, where the chemical formulae of rubber and gutta percha are shown to be different only, in that one is the cis and the other the trans form of poly-isoprene, it may be asked why the first is rubbery elastic at room temperature, while the other only possesses this property after being heated above 60° C. Gutta is crystalline at room temperature and melts at 60° C, proving that at this temperature the micro-Brownian motions of the chain elements become so intense, that the lattice forces are insufficient to keep these parts fixed.

In Fig. 8 some details of the structure of both materials mentioned are reproduced; the positions of the groups between which oscillations are possible, have been drawn with dotted lines. In both materials the CH_2-group 4b oscillates from one side of the isoprene unit to the other, being hindered by the CH_3-group 5. X-ray analysis has shown this group 5 to turn back a little further for rubber, the mean value of the angle 5 — 2 — 3 being 115°, whereas for gutta it is 125°. The oscillations of the CH_2-group 4b are still further facilitated in rubber because the vibrations of the CH_3-group 5 can have a greater amplitude; similar vibrations are impeded [2] in gutta by the CH_2-group.

It is assumed that at 60° C the heat vibrations of the CH_2-group 4 in gutta are sufficient to overcome these obstacles so that at this temperature gutta becomes comparable with rubber. The fact that chloroprene, having a similar structure to gutta (with, however, a Cl-atom instead of the CH_3-group, see p. 37), is rubbery at room temperature, is ascribed to the distance between the Cl-atom and the chain being a little larger (1.77 Å) than that between the CH_3-group and the chain (1.53 Å) and to the fact that the Cl-atom protrudes less from the molecule.

Fig. 9. Transformation interval for some linear polymers, expressed by the extension at the breaking point.

In Fig. 9 the transformation interval, expressed by means of the extension at the breaking point [3], is shown for two typical linear polymers, crude rubber (vulcanised rubber is discussed later) and polyvinyl chloride. The first material is a representative of the substances which show rubbery elasticity at room temperature,

---

[1] This problem has been thoroughly investigated by K. H. MEYER, *Die hochmolecularen Verbindungen*, Leipzig, 1940, p. 116. His considerations have been largely extended by C. W. BUNN, *Proc. Roy Soc. (London)*, A. 180 (1942) 40.

[2] This does not appear from Fig. 8, the diameters of the atoms having been designed on too small a scale.

[3] Data for raw rubber from E. A. HAUSER, P. ROSBAND und E. SCHMID, *Z. techn. Phys.* 9 (1928) 98; for vulcanised rubber from T. FUGIWARA and T. TANAKA, *Rubber Chem. Technol.*, 7 (1934) 610; for polyvinyl chloride: W. BUCHMANN, *Forschung*, 12 (1941) 174.

being therefore in every-day parlance indicated as "a rubber". Polyvinyl chloride, however, is hard at room temperature, only becoming rubbery on heating above say 80° C and is therefore usually indicated as "a resin".

The transformation interval covers a large range for both polymers. By vulcanisation of the rubber it shifts to higher temperatures, a point to which we will return on p. 674.

2°. Thermal recovery

Linear polymers often show very pronounced latent elasticity (see p. 666) in their transformation interval. In Fig. 10 it is demonstrated[1] that the lower the deformation temperature the more complete is the thermal recovery, the deformation being more plastic at higher temperatures.

a) Recovery temperature 85° C.   b) Recovery temperature 160° C.
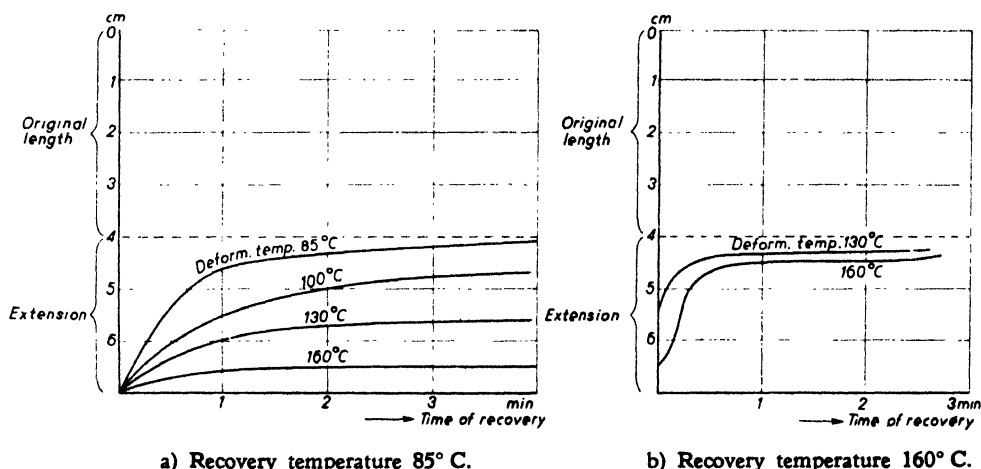
Fig. 10.   Thermal recovery of stretched pieces of polyvinyl chloride.

On the other hand a comparison of Fig. 10a and Fig. 10b shows that the thermal recovery is more complete for a higher recovery temperature, the viscosity then being lower.

3°. Tensile strength

From a technical point of view the relation between the mechanical strength and the molecular weight is of great importance. In Fig. 11 it is shown that for a natural[2] and for a synthetic polymer[3] an increase is observed on polymerisation. This is understandable, since the surface of cohesion between the molecules increases with their length. The fact that a maximum is reached can be explained by assuming that, at the corresponding polymerisation degree, the total secondary energy of cohesion between the chains becomes of the same magnitude as the primary energy

---

[1] W. BUCHMANN, Forschung, 12 (1941) 174.
[2] W. RÖHRS, H. STAUDINGER, R. VIEWEG, Fortschritte der Chemie, etc. München, 1939, p. 16.
[3] G. O. CURME and S. D. DOUGLAS, Ind. Eng. Chem., 28 (1936) 1123; S. D. DOUGLAS and W. N. STOOPS, Ind. Eng. Chem., 28 (1936) 1152.

connecting the units *in* a chain, leading to a rupture of the chain molecules themselves. Following this line of reasoning it is clear that for side groups with a high energy of cohesion, a strong product will be obtained, even if the length of the molecule is not very great.

On these grounds the great strength of the newly developed superpolymides (Nylon, Perlon) may be connected with the strong NH .... OC bonds between the chains, which are regularly ordered, so that the crystallisation on orientation is favourable.
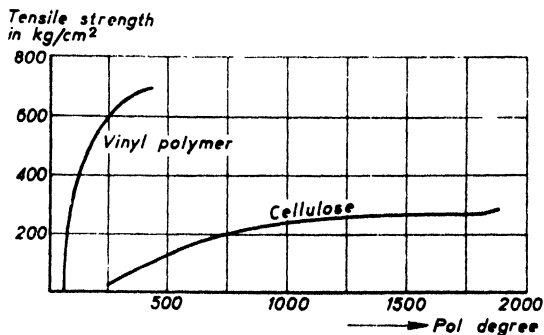
Fig. 11.   Influence of chain length on the strength of polymers.

c.   Influence of orientation

In the isogel state, deformation is so easy that, when stresses are applied in one direction to a linear polymer, the molecules are oriented. This leads to anisotropy, which can be controlled mechanically, and optically by X-rays. The stress-strain diagrams for the direction of orientation and that at right angles to it are different, as is seen in Fig. 12 for a sample of cellophane (cellulose xanthogenate) foil. The curve A refers to the direction parallel to the orientation and corresponds to a much stiffer material than curve B, referring to the direction at right angles to the orientation.
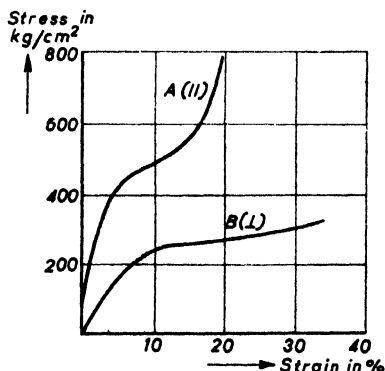
Fig. 12.   Stress-strain diagrams for oriented. cellophane in the two main directions.

The reason for this difference is that in the direction of orientation the molecules have been brought into better contact, cohering over a greater surface. This leads to important technical consequences, because in producing films some orientation very often arises in the process, leading to an unhomogeneous product. This has been demonstrated visually with the aid of strips of polystyrene[1] from such a film. Cut in the machine direction and loaded (with 78 g) at 84 °C, it shrank together (Fig. 13b), notwithstanding the fact that it was loaded, thus showing a thermal recovery, proving the presence of internal tensions as a consequence of the machine process.

At right angles to the machine direction however, the extension shown in Fig. 13c was obtained under the same conditions, indicating that plastic deformation was easily produced because no internal tensions were counterbalancing the applied stress.

---

[1] G. VAN ITERSON JR. and K. E. C. BUYN, *Kolloid-Z.*, 85 (1938) 60.

The changes in the molecular orientation during these tests can be expressed by means of optical anisotropy. The path-difference between the ordinary and the extraordinary ray of polarised light was 53 Å per mm thickness of the original piece, showing a noticeable birefrigence as a consequence of the machining process. For the piece illustrated by Fig. 13b this value was reduced to 13 Å, but in the case of Fig. 13c it increased further to 79 Å.

Another consequence of this anisotropy is that the tensile strength is greatest in the direction of orientation; in Fig. 12 it is 750 kg/cm^2, against only 300 kg/cm^2 in the right-angular direction. Important technical applications have been based upon this knowledge; in the rayon industry for example, the degree of orientation is one of the important means of controlling the properties of the finished product.

It is a limiting factor that a certain degree of orientation should not be surpassed, because otherwise the material would become too stiff, leading to an insufficient creasability.
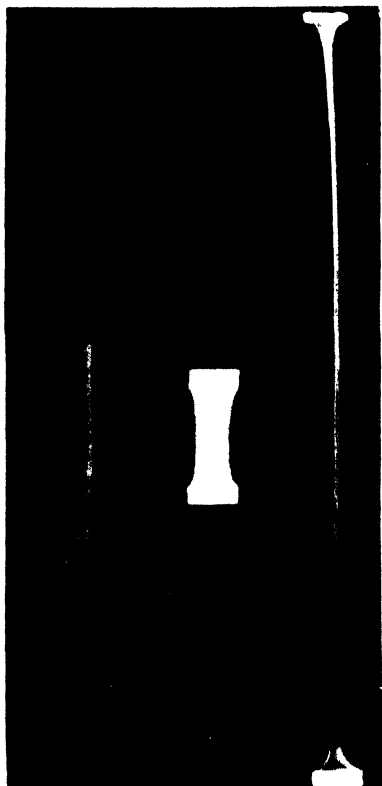
Fig. 13. Extension and thermal recovery of polystyrene, parallel and at right angles to the machine direction.
a) Unloaded film, cut in the machine direction. (Cross section 18 mm^2).
b) Loaded in the machine direction with 78 g at 84 °C.
c) Loaded at right angles to the machine direction with 78 g. at 84° C.

d. Crystallisation phenomena [1]

For certain chain polymers (e.g., cellulose), crystallised regions (so-called crystallites) are already present in the undeformed virgin state (see p. 619), but in other cases (e.g., rubber) they are formed during the stretching process. The molecules then become more and more oriented and are gradually forced into a crystalline lattice over certain parts of their length, which can be detected by means of X-rays (Fig. 14).

Most rubbers however, can also crystallise without being stretched, proving that the macromolecules are insufficiently mobile to snap into the crystalline lattice of their own accord, if time is available.

[1] For a general treatment see: L. A. Wood, *Advances in Colloid Science*, Vol II, New York, (1946) p. 57.
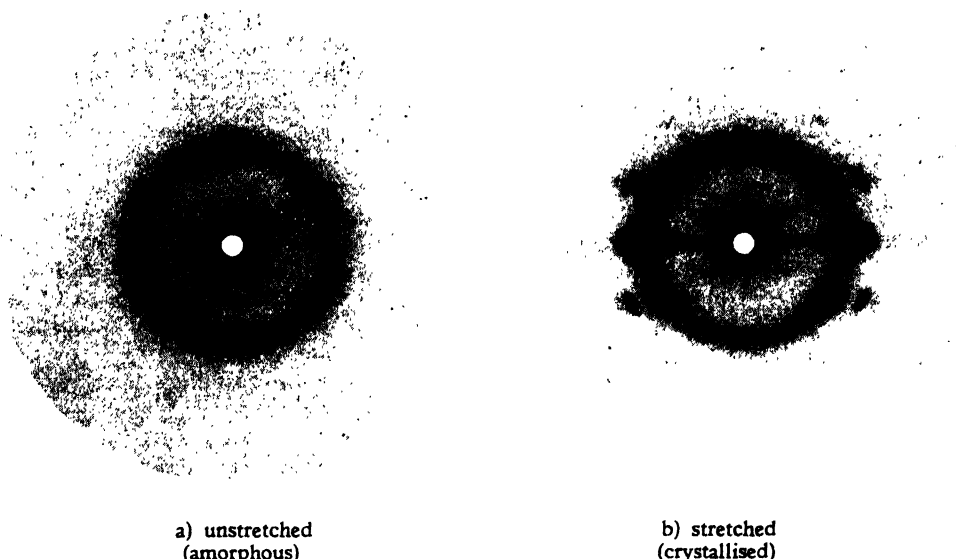
a) unstretched   b) stretched
(amorphous)   (crystallised)
Fig. 14.   X-ray diagrams of natural rubber.

This is illustrated in Fig. 15, where the degree of crystallinity is expressed by the relative volume. It appears, that when super-cooling rubber so quickly that it remains amorphous, a greater volume is found than for slow cooling, the latter leading to a more complete crystallisation. The location of T_m is not influenced by the crystallisation. The temperature at which crystallisation begins (T_k is at $+10°C$), is a kind of transition point, not identical with the brittle point T_b, the latter being lower. Perhaps it is identical with T_f the transition point gel $\rightarrow$ fluid, but this cannot be said with certainty as yet. These examples show how complicated the situation is, compared with low molecular substances.
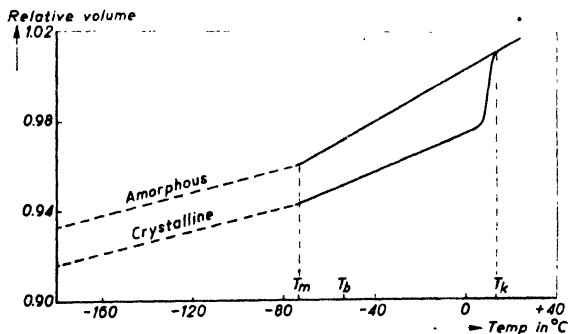
Fig. 15.   Crystallisation of rubber on cooling [1]

It has already been stressed above that in polymers the crystallisation phenomena will be of a different character than in a substance like NaCl; the difference is in their velocity. In NaCl all crystallising particles can find their place in the lattice immediately, being present in mobile surroundings (the solution). With polymers, the viscosity is roughly 10^8 times higher and so it takes a considerable time for the long chains to find sufficient

[1] N. BEKKEDAHL, j. Research Nat. Bur. Standards, 13 (1934) 411.

potential energy minima along considerable parts of their surface for crystallites to form. Fig. 16 shows this gradual crystallisation at various temperatures, expressed by means of the volume decrease; it is quite clear that the speed of crystallisation is strongly dependent on temperature. The greatest speed is reported to be at —25 °C for natural rubber; the maximum and minimum temperatures between which crystallisation can occur are + 15° C and —50° C respectively.

For neoprene the temperature range is between —35 and +32° C. Polyethylene is one of the synthetic polymers with the strongest crystallisation tendency; from 55 to 75% crystalline at room temperature is no exception. We return to this important point on p. 669.

The size of the crystallites can be of the order of $500 \times 500 \times 200$ Å. Upon heating, the crystals do not melt at a single temperature, as is shown in Fig. 17. As the melting point depends on the temperature at which the crystals have been formed (Fig. 17), the melting range is explained by assuming the co-existence of crystals of several types.

Fig. 16. Crystallisation of natural rubber measured by the volume decrease.

A proof of this assumption is that according to Fig. 17 for instance, the crystals formed at —30° C are completely molten at 0° C whereas on the other hand, crystals formed at 0° C, can withstand a temperature of 15° C.

The crystallisation is reflected in values of the specific volume, in X-ray diffraction, in optical double refraction, in an increase of the modulus of elasticity and also in the flow properties; as is shown in Fig. 18.

If no complications were involved, one would expect the plastic flow to increase continuously with the degree of extension, but for crude rubber it appears[1] that the flow decreases after surpassing a limit of about 450% elongation. This can be ex-

Fig. 17. The melting range of rubber crystals as a function of the crystallisation temperature.

plained by assuming that, above this limit, crystallisation results in the chains fitting so strongly in the crystalline lattice that plastic flow is more and more suppressed.

Here, the very important bearing of crystallisation phenomena on the technical value of rubbers must be referred to. Those rubbers which do not crystallise on

---

[1] L. R. G. TRELOAR, *Trans. Faraday Soc.*, 36 (1940) 538.

stretching (GR-S, the co-polymer of butadiene and styrene) have a low strength unless they are mixed with carbon black. Obviously, the snapping of the chain molecules into the lattice is necessary to make the chains cohere sufficiently. On the other hand, the percentage of crystallites should not be too large, otherwise the material becomes hard and loses its rubbery properties. Polyethylene, which in the unstretched condition is from 55 to 75% crystalline, surpasses the tolerable limit. The fair average seems to be about 30%, which can be concluded from the unique mechanical properties of natural rubber, which crystallises to this extent[1] on stretching.

The fact that carbon black brings the mechanical properties of non-crystallising materials to a level comparable with those of natural rubber, can be explained by the positive heat of wetting. This proves that the carbon black adheres strongly to the rubber molecules, "glueing" them together.

The reason that GR-S does not crystallise can be found in the various possibilities of mutual rearrangement of the monomers in the chain during the polymerisation. They can attach themselves, not only head to tail, but

Fig. 18. Plastic flow of crude rubber as a function of elongation (50° C).

also in the head to head position. Furthermore, there are possibilities of cis-trans isomerism, and 1.2 and 1.4 polymerisation of the butadiene. Last, but not least, the butadiene and styrene groups may be arranged in an irregular order. Altogether, there are so many factors leading to the formation of irregular chains that it is not surprising that crystallisation hardly takes place in GR-S.

e. Internal and external plastification

One of the most common ways of influencing the secondary coherence of macromolecules is by adding a lubricant, a so-called plasticiser. This must have a negative value of $\triangle$ F (equation 1 on p. 154) with regard to the polymer molecules in order to penetrate between them. Its own molecules however, must have a very weak mutual coherence, making gliding easily possible; as a rule therefore it is a liquid or an easily deformable solid (waxes, fats).

There is also, however, another way of weakening the coherence between macromolecules, namely by producing an adequate amount of co-polymer (see p. 26). Indicating the structural units of the original polymer by A and of the subsequently introduced chain elements by B, the latter must have a secondary coherence with regard to themselves, or with regard to the A-groups, which is weaker than that between the A-groups themselves.

If, for instance, the monomer of the type A is a solid but that of the type B a liquid (or still better a gas), the coherence between the co-polymer molecules AB

[1] A. J. WILDSCHUT, *Technol. and Phys. Investig. on Natural and Synthetic Rubbers*, Amsterdam 1946.

will be weak at those points where two B-groups are in contact with each other and perhaps also[1] at the points where B-groups are in contact with A-groups. The more B is introduced, the weaker will be the coherence of the molecules, and by this means the properties of the material may be controlled. This has become of practical importance in developing rubbers in which long chains with a loose coherence are necessary.

These considerations can be followed taking polystyrene[2] as an example. In Fig. 19, the transition points have been expressed by means of the volume changes.
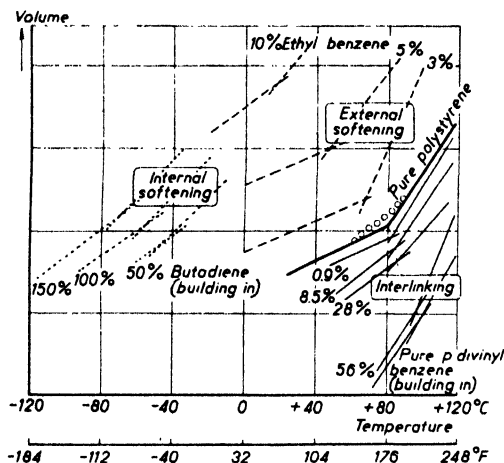
Fig. 19. Decreasing and increasing the secondary coherence of polystyrene molecules.

For pure polystyrene a transition is observed at about 80° C; this will be indicated as the transformation point although actually it is a large region (indicated by 00000). Above this temperature the material is a fluid; below 80 °C it is a solid.

On adding ethylbenzene ($CH_3—CH_2—C_6H_5$), the softening action is immediately demonstrated by a shifting of the transformation point to lower temperatures; a 10% addition reduces the transformation point from 80 to 20° C. This is of technical interest, because rubbery properties are thus obtained at room temperature. A much greater effect can however be produced by making a co-polymer with butadiene ($CH_2 = CH — CH = CH_2$), a gas with a boiling point of —3 °C. A 50% addition of butadiene reduces the transformation point to —40° C, but the material obtained is still rather stiff at room temperature. Therefore in GR-S (see p. 38) about 60% of butadiene is introduced. The transformation interval of this material extends to about — 80° C, opening a wide field for technical applications as a synthetic rubber.

The opposite effect, namely that of stiffening a material by interlinking may be referred to here. In the case[3] of polystyrene this can be effected by means of p-divinyl benzene,

$$\begin{array}{c} CH = CH_2 \\ | \\ C_6H_4 \\ | \\ CH = CH_2 \end{array}$$

having two double bonds, leading to the inter-

linked product shown in Fig. 20.
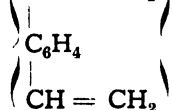
The result, a shifting of the transformation point to higher temperatures, can be observed in Fig. 19. The same type of effect will be encountered, when dealing with the vulcanisation of rubber (p. 674).

[2] This depends on whether the liquid (or the gas) B is strongly absorbed by A.

[1] Data borrowed from K. UEBERREITER, Angew. Chem., 53 (1930) 247; Z. phys. Chem., B 45 (1940) 361; B 46 (1940) 157; E. JENCKEL, Kolloid-Z., 100 (1942) 163.

[2] E. L. KROPA and T. F. BRADLEY describe other interlinking actions, namely of glycol maleate resins on vinyl acetate, methyl methacrylate and polystyrene, Ind. Eng. Chem., 31 (1939) 1512.

$\cdots\cdots\text{—CH—CH}_2\text{—CH—CH}_2\text{—}\left(\text{CH- -CH}_2\right)\text{—CH—CH}_2\text{—CH--CH}_2\text{—CH—CH}_2\text{—}\cdots\cdots$
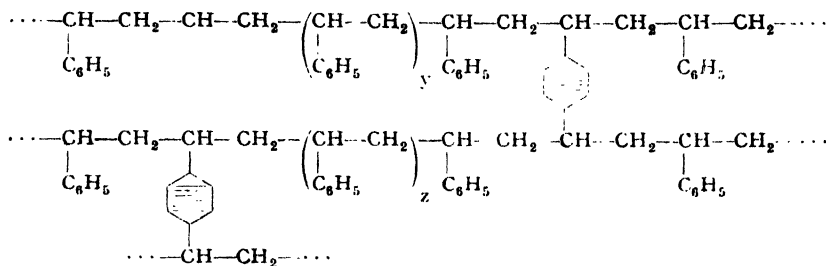
Fig. 20. Polystyrene, interlinked by means of p-divinyl benzene.

f. Influence of the molecular weight

Experimental evidence is available to show that, above a polymerisation degree of about 1000, T_b and T_m become practically independent of molecular weight. It is a remarkable observation that T_b decreases, whereas T_m increases, with the degree of polymerisation. The explanation of these phenomena is difficult at present and is considered to be outside the scope of this book.

§ 4. THREE-DIMENSIONAL POLYMERS

a. The transformation interval

With some necessary changes similar considerations as encountered on p. 660 for linear polymers, can be given for the growth of globular macro molecules, leading to the picture[1] given in Fig. 21. The sol in Fig. 21a may contain two types of particles. First, small freely rotating globules, to be denoted as α-particles, which are "liquid" and soluble[2] as well. Secondly, it contains so-called β-particles, built up from α-particles by condensation to such a size that they no longer exhibit Brownian motion, being therefore "solid". They are however coherent to each other and to

a) Sol   b) Gel (soluble and fusible)   c) Gel (partly soluble but infusible)   d) Hard solid system (insoluble and infusible)

◌ α-particle Freely rotating; "liquid and soluble"

◯ β-particle. Not rotating; "soluble"

⬤ γ-particle Not rotating; "solid and insoluble"
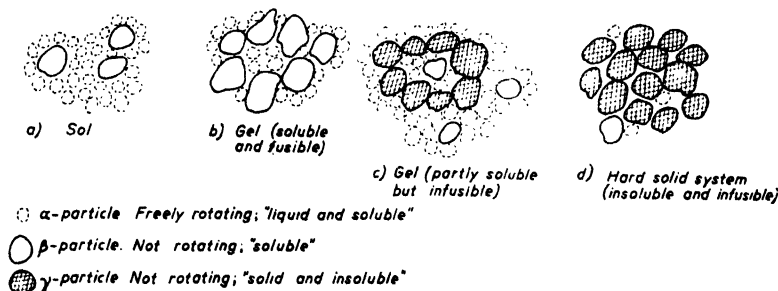
Fig. 21. Changes in colloidal structure on three-dimensional polymerisation.

[1] R. HOUWINK, Elasticity, Plasticity and Structure of Matter, Cambridge, 1937.
[2] In solvents, in which the monomer is soluble.

the α-particles by means of secondary bonds, thus still being soluble. Cooling such a sol will increase the number and the size of the β-particles, until they cohere throughout the mass, leading to a gel of the still soluble and fusible type (Fig. 21b) discussed already.

If the sol of Fig. 21a is not solidified by cooling, but by polymerisation, the situation shown in Fig. 21c arises where a new type of particle, the γ-particle, must be introduced. By this is meant a β-particle which has grown so much that it is attached to others by primary bonds at least at one spot, so that they can no longer be split off by solvents. This leads to another type of gel, namely that which is still partly soluble but infusible.

In the hard solid state of Fig. 21d a few α- and β-particles are assumed to be still present, it being practically impossible to bring condensation[1] reactions to an end completely (see p. 41). Their number is so small however, that the material is practically completely insoluble and infusible.

b. Plastic and elastic properties

For the present section, the gel-state of Fig. 21c is of special interest. The sol exhibits a purely viscous flow, but when the gel-state is introduced by subsequent polymerisation, the flow curve gradually approaches the form shown by KRUYT in Volume I, related to quasi-viscous flow. This is a consequence of the fact that the net must now be broken by the shearing forces; this type of deformation is therefore accompanied by elastic (namely, highly-elastic) phenomena. These will become still more important on subsequent polymerisation to the final stage, a sufficient mobility of the net elements then being impossible. A certain region of polymerisation will thus exist in which rubbery elasticity is observed, corresponding to the structure shown in Fig. 21c.

Examples of this can be found with phenol-formaldehyde, urea-formaldehyde, glyptal, and shellac resins (for formulae see p. 39). In Fig. 22 the change of highly elastic deformability is demonstrated[2] for a

Fig. 22. Change of high-elasticity and solubility for some three-dimensionally hardening resins during polymerisation at 100—125° C.

cresol-formaldehyde[3] resin and for shellac, showing that in both cases a certain degree of polymerisation is necessary for high-elasticity to be observed (the sol must

---

[1] Although polymerisations may not be excluded, they will not be discussed here, most globular polymers being formed by condensation processes.

[2] R. HOUWINK, *Trans. Faraday Soc.*, 32 (1936) 111.

[3] Containing 58—60% m-cresol.

first be converted into a gel). Then it increases very rapidly, reaching values of the order of 200%, thus entering into the region of rubbery elasticity. After reaching its maximum a quick drop is observed, connected with the transition into a one-phase system. The change of solubility is represented by dotted lines and from it, it may be inferred that the region in which high-elasticity is observed corresponds to a percentage of soluble constituents roughly between 25 and 60% for the cresol resin and between 25 and 95% for shellac.

The fact that the whole process is executed in a much shorter time for the cresol resin is due to its greater reaction velocity. One of the important conclusions from these experiments is that *rubbery elasticity*[1] *is not restricted to linear polymers, but can also be observed with globular macromolecular substances.*

As soon as these resins have become one-phase systems they behave as hard rigid materials like crystals. On extension, HOOKE's law is followed and the modulus of elasticity is of the same order (50,000 kg/cm²) as that of crystals, which is in accordance with the fact that in both cases primary bonds have to be deformed. We will not enter into the problems connected with this stage[2], the dimensions of the kinetic units now surpassing the limits of the colloidal region (cf. p. 24). We will merely point to the fact that the tensile strength of the hardened resins is about 500 to 1000 times lower than might be expected on theoretical grounds as has already been mentioned for glass, on p. 657. This fact may be referred to briefly here, because it is connected with the colloidal structure in the transformation interval. On similar grounds as those discussed on p. 657 it may be assumed that at the moment of mutual contact between the colloidal aggregates, there are steric reasons for all theoretically possible bridges not actually being formed, leading to chinks in the material.

§ 5. INTERLINKED SYSTEMS

a. General considerations

Above it was remarked that the interlinked systems form, to a certain extent, a transition between the linear and the globular polymers. On these grounds their properties may be expected to be somewhere between those of the two systems mentioned, and this can actually be observed.

From a practical point of view, the interlinking processes of macromolecules are of greatest importance, being applied on a large scale in practice and in nature also. Restricting ourselves to a few examples, we will mention only the vulcanisation of rubber and the bridge formation between oil molecules in paints by means of resins or oxygen. The analogy with three-dimensional polymerisation is particularly striking in the latter case, as will be seen from Fig. 23.

Drying oils contain glycerides with three fatty acid residues, having the following steric form and as was shown on p. 36, fatty acids have the tendency to become interlinked.
This takes place on the three arms, pictured in Fig. 24, and leads to an intensive net formation of the three-dimensional type.

---

[1] A different question is whether this is ideal high-elasticity (see p. 679). Presumably the necessary conditions are not fulfilled.

[2] See R. HOUWINK, *Rubbers and Resins, their Structure and Properties*, Elsevier, Amsterdam, (in the press).

The results are well known. On drying, paint gradually thickens, then becomes highly elastic, remaining for years in this state, the useful period of its existence, finally becoming hard and brittle. From this it can be concluded that the slow reaction
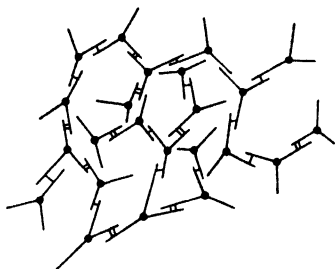
Fig. 23.   Net formation in oil films.
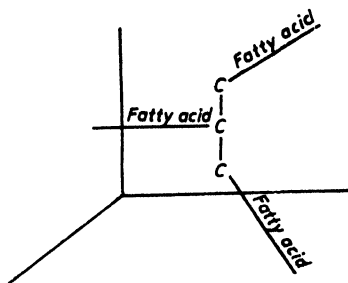
Fig. 24.   Three molecules of fatty acid, attached to a (central) glycerine molecule.

rate, maintaining the gel state for years, is one of the striking features of paints. If the reaction rate were greater, the life time of an (elastic) paint would be seriously restricted. When the hard one-phase system has been formed, the technical

Fig. 25.
Influence on T_m and T_b by vulcanising rubber.

value of the film ceases to exist, cracks easily being formed by expansion and contraction of the coated surface (swelling by moisture and expansion by the heat of the sun).

The influence of the degree of interlinking on T_m has been measured by BEKKEDAHL[1] and in Fig. 25 his results are shown, together with some results on the change of T_b.

As expected both transition points move to higher temperatures, which is in accordance with the results of interlinking polystyrene by means of p-divinyl benzene (p. 670).

b.   Behaviour on deformation of interlinked systems

1°.   The stress-strain diagram

The changes in the stress-strain diagram by interlinking are shown in Fig. 25[A] for the case of vulcanising rubber. The stiffening of the material is demonstrated clearly, still greater stresses being necessary for a certain extension. Generally speaking however, the modulus of elasticity (a measure for the stiffness) is very low. As a rule it is about 20 kg/cm² for rubbery substances whereas for hard rigid materials like

[1] For a general survey see: R. F. BOYER and R. S. SPENCER, Advances in Colloid Science, Vol 2, New York, 1946.

crystals, hardened resins (and also for completely stretched crystallised rubber) it is of the order of 10^5 kg/cm^2. This indicates that the stretching of amorphous rubber causes molecular phenomena quite different from the stretching of the harder materials. In the case of rubber it is the orientation which must be carried out, while in the case of the harder materials it is the stretching of primary bonds.

Returning to Fig. 25 A it may be remarked that at the beginning, the stiffening by vulcanisation (up to 5% of bound sulphur) is accompanied by an increase in tensile strength and both observations can be understood on a basis of the reinforcement of the mutual coherence between the chain molecules.

Fig. 25 A.   Influence of vulcanisation on the stress-strain diagram of rubber.

Now however, the question arises as to why the strength goes through a maximum when 5% of sulphur has been bound. This can be explained by pointing to the favourable influence of crystallisation, already mentioned on p. 666. X-ray measurements have shown[1] that on proceeding with the vulcanisation the crystallising power also passes through a maximum, corresponding to a percentage of bound sulphur similar to that necessary to reach the maximum tensile strength.

Fig. 26.   Theoretical and experimental stress-strain curve for rubber

Of fundamental importance is the fact that it has proved possible to calculate the stress-strain diagram for a material like rubber starting from some simple assumptions concerning the free rotation, and flexibility of the coiled molecules. Thus, GUTH and JAMES[2] derived the formula:

$$\tau = KT \ (L - 1/L^2) \qquad (3)$$

where τ = tension, L = final length

[1] J. E. FIELD, *J. Applied Phys.*, 12 (1941) 23.

[2] For an extensive summary see: E. GUTH, H. M. JAMES and H. MARK, *Advances in Colloid Science*, Vol. 2, New York, 1946.

and Fig. 26 shows that this result corresponds to curves obtained in practice up to an elongation of about 350%. This is the limit above which the assumptions about the statistical coils are no more valid. On the same grounds formula (4) was calculated for expressing the differential modulus of elasticity E.

$$E = KT (1 + 2/L^3) \tag{4}$$

2°. High-elasticity[1]

Rubbery elasticity can be observed in its most ideal form in slightly interlinked polymers. On the one hand, the molecules are large enough to create the conditions necessary for high-elastic deformations; on the other hand, the bridges prevent the molecules from gliding over their neighbours, thus suppressing plastic deformations. As was pointed out above, it is only necessary that the mutual coherence between the chains is weak enough for the chain elements to rotate easily, and that the degree of interlinking is kept below certain limits, since otherwise the material becomes too hard.

In this paragraph these problems may be considered a little more deeply. Theoretical considerations lead to the conclusion that single bonds of the type $C - C$, can rotate freely about their axes. This is no longer true in $CH_3 - CH_3$, due to a steric interaction of the H-atoms, leading to an energy barrier of 3 kcal per mol. By spacing these CH_3-groups farther apart, as in $CH_3 - CH = CH - CH_3$, their mutual interaction diminishes, leading to a barrier of only 0.5 kcal per mol. This presumably is also one of the causes why polyethylene is not an ideal rubber. The double bond in natural rubber seems to serve a twofold purpose: firstly, improving the rotation possibilities and secondly, making vulcanisation possible.

Fig. 27. On heating extended rubber, the stress increases; at low extensions, a linear relation even exists.

In part I (Kruyt) the high-elastic phenomena were subdivided into *potential* and *entropic elasticity* and we will now consider these two kinds more indetail[2] (see also this book, the chapter by J. J. HERMANS).

In a metal spring the energy added on extension is stored as potential energy, removing the atoms from the bottom of their potential troughs. The thermodynamic balance is maintained by a drop of temperature, heat being given off to the surroundings. On release, the atoms fall back into their potential energy

---

[1] For an extensive summary see: E. GUTH, H. M. JAMES and H. MARK, *Advances in Colloid Science*, Vol. 2, New York, 1946.

[2] Cf. for detailled treatments: W. B. WIEGAND and J. W. SNIJDER, *Trans. I.R.I.*, 10, (1934) 234; K. H. MEYER, *Chem. Revs*, 25 (1939) 137; E. WÖHLISCH, *J. prakt. Chem.*, 160 (1942) 217; R. HOUWINK, *Rubbers and Resins, their Structure and Properties*, Elsevier, Amsterdam (in the press).

minima, a process which is accompanied by a rise in temperature. It is in accordance with this, that on heating a piece of metal (stressed or unstressed), it tries to elongate itself further, the atoms having larger heat-vibrations, removing themselves further from the positions of minimum potential energy. The tension however of a stressed piece of metal decreases on heating, since part of the stored up potential energy is then liberated.

Completely different observations can be made for certain polymers, e.g., a piece of stretched rubber. Fig. 27 shows that for extensions up to about 350% the tension increases[1] on heating and that there even exists a proportionality[2] between tension and absolute temperature (above about 350% extension this proportionality vanishes). It has also been shown that heating leads to a contraction (a negative thermal expansion!) of the extended rubber. The thermodynamic balance demands that on extending rubber it will get warmer, and this is actually observed.

This peculiar behaviour of certain polymers can be understood with the aid of a thermodynamic analysis, starting from the fundamental equation:

$$dQ = T \, dS = dU - dA \tag{5}$$

where Q = heat exchange
U = internal energy (potential plus kinetic)
A = external energy (mechanical work)
S = entropy

Stretching a piece of rubber over a length dl by applying a tension τ, this equation can be written:

$$T dS = dU - \tau \, dl + p \, dv$$

As the volume change can be neglected[3], one may write

$$dU = T \, dS + \tau \, dl \tag{6}$$

The free energy F is given by

$$dF = d(U-TS) = dU - TdS - SdT$$

and introducing (6):

$$dF = \tau \, dl - SdT$$

dF is a total differential and so one can write:

$$\left(\frac{\partial \tau}{\partial T}\right)_l = -\left(\frac{\partial S}{\partial l}\right)_T \tag{7}$$

From (6) it can be further concluded that

$$\tau = \left(\frac{\partial U}{\partial l}\right)_T - T\left(\frac{\partial S}{\partial l}\right)_T \tag{8}$$

Introducing into this equation the experimental result that up to a certain extension

$$\tau = CT \quad (C \text{ is a constant}),$$

it is concluded that:

$$CT = \left(\frac{\partial U}{\partial l}\right)_T + T\left(\frac{\delta \tau}{\delta T}\right)_l \tag{9}$$

As $\tau = CT$, is also true: $\left(\frac{\delta \tau}{\delta T}\right)_l = C$

---

[1] Except at very low temperatures, see later.
[2] Combined with the fact, that the τ-T lines meet at the origin, so that for $T = 0$ also $\tau = 0$.
[3] Rubber is practically incompressible.

Then it follows from (9) that

$$CT = \left(\frac{\partial U}{\partial l}\right)_T + CT, \text{ and thus: } \left(\frac{\partial U}{\partial l}\right)_T = 0$$

This would lead to the conclusion that in the region of extension where τ increases proportionally with T, the *internal energy is not changed during an isothermal extension*. The only way therefore to maintain the thermodynamic balance is for the work (dA), done on the test piece, to be given off in the form of heat (—dQ). As —dQ = —TdS, the entropy decreases. But, the entropy always tends towards a maximum and thus the rubber will try to absorb heat from its surrounding atmosphere, which however must be accompanied by a contraction[1]. In such a case the entropy appears to be the moving power for the contraction, leading to the above mentioned expression: *entropic elasticity*. This can be further understood with the aid of the Boltzmann formula, connecting the entropy with the probability W_1 of the molecules being in the most random situation with regard to each other:

$$S_1 = k \ln W_1 \tag{10}$$

On stretching, the molecules are brought into a state of equal internal energy (only if $\tau = CT$) but of a better order (orientation), corresponding to a smaller probability W_2. The free energy of this state, and consequently the elastic tension, is measured by the entropy difference.

$$S_1 - S_2 = k \ln \frac{W_1}{W_2} \tag{11}$$

In terms of molecular kinetics this can be expressed as follows. In the unordered state the molecules are hammering at each other equally in all directions, as a consequence of thermal motion. After orientation however, all directions are no longer equal and therefore the hammering also shows a certain preference. It will be strongest at right angles to the direction of orientation, the bonds in that direction being weaker. The molecules will try to obey this unidirectional surplus of impulses and will strive after the maximum of disorder already mentioned, corresponding to the maximum of entropy.

The metal spring and rubber may be very sharply contrasted by means of the equation (8):

$$\tau = \left(\frac{\partial U}{\partial l}\right)_T - T\left(\frac{\partial S}{\partial l}\right)_T$$

for the rubber: $\left(\frac{\partial U}{\partial l}\right)_T = 0 \quad \tau = -T\left(\frac{\partial S}{\partial l}\right)_T$

for the spring: $\left(\frac{\partial S}{\partial l}\right)_T = 0 \quad \tau = \left(\frac{\partial U}{\partial l}\right)_T$

This clearly expresses that for rubber the increase of tension on isothermal stretching is an entropy effect, whereas for the spring it is of a potential energy nature.

If the extension is carried out *adiabatically* (which is practically the case when stretching rubber very quickly), T is not constant and thus the internal energy increases. The work done is then transformed into kinetic energy (heat), in contrast

---

[1] On this principle a heat-engine has been constructed; cf. W. B. WIEGAND and C. W. SNIJDER, *Trans. I.R.I.*, 10 (1934) 234.

to the metal spring, in which it was stored as potential energy. From this point of view one can also talk about the *kinetic elasticity* of rubber.

Considering further the conception of an extension without change of internal energy, this would make rubber comparable with an ideal gas. It would mean that during stretching to about 350% of the initial length, no bonds (either primary or secondary) would be deformed and no rotations (where there is an energy barrier to be overcome) would occur. These conditions, which may be realised in an ideal gas, are not fulfilled in a real gas and to an even smaller extent in a compact substance like rubber; it has been shown that a potential barrier of about 0.5 kcal per mol has to be reckoned with. In our opinion [1] the solution of this problem can be found by assuming that there are not only positive changes of internal energy (deformation of valences, rotations) but negative ones as well (pressing of chains against each other, crystallisation), neutralising each other exactly in the range of extension up to 350%. This range is denoted as that of *ideal high-elasticity* and its extent in many respects is a measure for the quality of a rubber. The range is rather wide for natural rubber and neoprene; it only exists to a very limited extent in strongly polar materials like proteins[2].

The trend of the curve for 500% extension in Fig. 27, where no proportionality between τ and T is observed, may be explained by supposing that here the orientation of the molecules has become so pronounced that the positive and negative effects do not exactly compensate each other any longer, leading to a surplus of one or other of them.

Fig. 28. Splitting the stress-strain curves of some polymers into kinetic and potential parts.

It has proved possible [3] to make a thermodynamic analysis of the complete stress-strain diagram, splitting it into kinetic and potential parts. This has been done in Fig. 28, showing that ideal elasticity $\left[\left(\dfrac{\partial U}{\partial l}\right)_T = 0\right]$ is not [4] observed for the rubbers of the perbunan type, as has

---

[1] R. HOUWINK, *Z. phys. Chem.*, A. 183 (1938) 209.

[2] H. J. WOODS, *Nature*, 157 (1946) 229.

[3] A. J. WILDSCHUT, *Technol. Physical Investigations on Natural and synthetic Rubbers*, Amsterdam 1946. At a certain tension $\left(\dfrac{\partial \tau}{\partial T}\right)_l$ is measured by increasing the temperature and with the aid of (9) the term $\left(\dfrac{\delta U}{\delta l}\right)_T$ can then be calculated.

[4] Strictly speaking, this is not the case for natural rubber either, as appears from Fig. 28.

already been remarked. This may be due to the fact that for these polar rubbers[1], any mutual movement of the molecules will give rise to important changes of internal energy, too large to be counterbalanced by opposite effects. It appears moreover, that on stretching natural rubber above 350%, $\left(\dfrac{\partial U}{\partial l}\right)_T$ first becomes negative, later, however, becoming positive again. The negative contribution may be due to crystallisation effects, and the positive to the stretching of the crystals formed, which obviously eventually exceeds the negative crystallisation effect.

Returning now to the left part of the curve, for a 350% extension it has been shown that, at about —70° C, it suddenly turns in the same direction as for hard materials like metals. Referring to p. 661 it may be remembered that —70° C is the temperature T_b, at which rubber freezes, loosing its high-elastic properties. In accordance with the above developed theories, the conditions for ideal high-elasticity are now no longer present, because the molecules are not sufficiently movable. At about —70° C the kinetic (or entropic) elasticity is therefore transformed into potential elasticity.

---

[5] For other polar substances a region of ideal elasticity is observed, and has been experimentally verified for polyvinyl acetate and for the inorganic rubbers.

XIV. ASSOCIATION COLLOIDS

by H. L. Booij

Department of Medical Chemistry, University of Leyden.

§ 1. INTRODUCTION

During the controversy over the question whether polymeric substances, which show colloidal properties in solution, are to be considered as molecules or as particles built up from many small molecules, the question was raised whether soap solutions might not serve as the model for polymeric colloids. It was in fact known that soap solutions owe their colloidal properties to a reversible association of the fatty acid anions or molecules. It became however ever more clear that the polymeric colloids must be considered as macromolecules — that is to say many monomers are bound together by ordinary valencies to form a polymer. It can be deduced from many properties of the polymers that these are indeed macromolecules.

At a time when this controversy over the polymers had not yet been settled — for example it was for a long time asserted of cellulose that it was built up of small molecules held together by strong lattice forces — STAUDINGER divided organic colloids into:

A. Association colloids, in which the molecules are reversibly associated to a colloidal particle;

B. Eucolloids, in which the colloidal particle is the same as the molecule.

According to STAUDINGER[1] the association forces correspond more or less with crystal-forming forces. An important characteristic is that such associated substances will frequently make their appearance in another solvent as single small molecules.

We prefer the concept association colloids above other terms — for example colloidal electrolyte — because we want to leave open the possibility that non-electrolytes or undissociated electrolytes form particles with colloidal properties by association. The association of small molecules or ions into micelles seems to us the more characteristic, not the electrolyte character which naturally plays a great part in soap solutions. Indeed particles which owe their character as colloids to their large molecules (for example proteins) also come under the concept of colloidal electrolyte.

Intermolecular forces are also active in gases and solutions, forces which lead to the deviations from the laws of ideal gases and solutions and are recognisable in the constant a of VAN DER WAALS. When these forces are large, molecules will orient themselves in a particular way with respect to each other. Thus we shall be able to distinguish:

---

[1] H. STAUDINGER, *Z. angew. Chem.* 42 (1929) 37, 67.

a. Swarms, when the dipoles point somewhat towards each other.

b. Aggregates of varying composition.

c. Molecular complexes (double molecules or multiple molecules) which have a definite composition.

Since a colloidal particle is characterised exclusively by its size, it is our task to investigate what conditions a molecule must satisfy for a colloid to be produced by association. It does not seem to us particularly probable that small dipoles in an aqueous medium will be able to form large particles by association.

In general they will be able to unite into small particles (see Fig. 1), long structures (b) being less stable than ring-shaped structures (c). It can be predicted theoretically that molecules with a dipole and appreciable non-polar parts in the molecule can lead to stable formations (d).

If the medium is not water but a dipoleless liquid, then the conditions will naturally be quite different. Thus for example, tetra-pentyl-ammonium iodide ($N(C_5H_{11})_4$ I) is dissociated into ions in water and acetone, is polymolecular in chloroform while it forms a colloidal solution in carbon tetrachloride.

Fig. 1.   Association of dipoles (according to A. VON BUZAGH).

We can provisionally declare that there is a good chance of association — in water as solvent — when the substance is a polar compound with a large dipole moment and anisodimensional structure. It may be expected that the tendency to association will become greater with an increasing number of carbon atoms. It is obvious that this association can be influenced by many factors: the dipole properties of the molecule, the dielectric constant of the medium, the steric structure of the molecule, the temperature, the concentration.

With anisodimensional molecules states have indeed been observed which lie between the crystalline and the amorphous (mesophases, liquid crystals). MARK[1] recognizes the following transitions between solid crystals and amorphous liquids:

1°. Three-dimensional crystal: the centres of gravity of the units are fixed (apart from vibration), rotations are not possible. Examples: hexamethylene tetramine, urea.

2°. Crystal with rotating molecules: the centres of the particles are fixed; rotation about one or more axes is possible. Examples: NH_4Cl, sodium stearate at higher temperatures.

3°. Smectic state: the centres of gravity of the cells are mobile in one direction, rotation about one axis is permitted. Examples: sodium oleate; phenetole-azoxybenzoic acid allyl ester at 72°.

4°. Nematic state: the centres of gravity of the cells are mobile in two spatial directions; rotation about one axis is permitted. Example: phenetoleazoxybenzoic acid allyl ester at 88°.

---

[1] H. MARK, *The general chemistry of high polymeric substances*, (1940) p. 199.

5°. Amorphous liquids: the centres of gravity of the cells are mobile in three directions
in space; rotation about three axes perpendicular to each other is possible.
Examples: CCl_4, Hg.

  As much as fifty years ago KRAFFT [1] observed that the gelatination temperatures
in soaps correspond fairly accurately with the melting points of the free fatty acids.
Approximately the same forces have therefore to be overcome in the melting of the
gels and of the free fatty acids. If we agree that many soaps occur as mesophases,
we may suspect that mesophases may also be present in soap solutions.

§ 2. ASSOCIATION IN SOAP SOLUTIONS

  A clear, not too concentrated soap solution shows no TYNDALL effect, no streaming
double refraction, while no particles can be observed with the ultramicroscope.
Nevertheless almost all the physical properties of the soap solutions change at a
particular concentration and these changes point to aggregates of molecules or ions.

  Thus KRAFFT [1] found a much too high molecular weight in many cases in measure-
ments of the rise of the boiling point of the homologous series of salts of the fatty
acids. The molecular weight begins to show a discrepancy at the nonylate and the
discrepancy continues to increase with increasing carbon chains. DONNAN and POTTS[2]
attempted to probe into the basis of the emulsifying power of soaps and for this
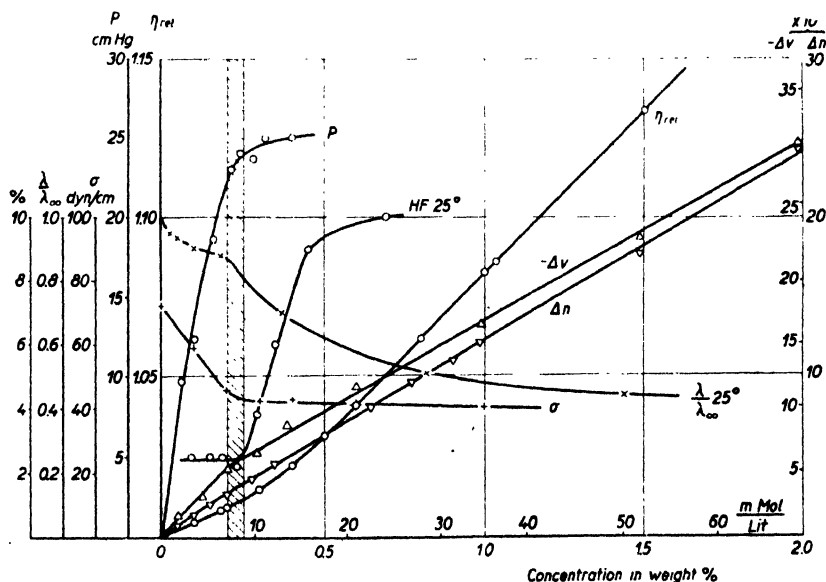
Fig. 2. Influence of the formation of micelles on the properties of Na dodecylsulphate at 20° and
25° C. (HF = conductivity for high frequency current, P = osmotic pressure, γ/γ_{∞} = conductivity
coefficient, σ = surface tension, η_{rel} = relative viscosity, IV = change of specific volume, Δn =
change of refractive index). From HESS, PHILIPPOFF, and KIESSIG, 1939.

[1] F. KRAFFT, Ber., 29 (1896) 1328.
[2] F. G. DONNAN and H. E. POTTS, Kolloid-Z., 6 (1910) 208.

purpose measured the interfacial tension between a hydrocarbon and solutions of the salts of fatty acids. They observed a clear lowering of the interfacial tension just at lauric acid. These changes can most readily be explained by an association into larger aggregates and it is clear that all properties, which are connected with the number of particles in the solution, will be affected by this. The osmotic properties — measurable from the lowering of the freezing point, rise of boiling point, vapour pressure, displacement of the dew point — were the subject of studies by Mc BAIN [1]. For each soap a definite concentration (the critical concentration) can be indicated at which these changes set in. The solubilisation of organic substances insoluble in water begins at the critical concentration and afterwards rises steeply. If one plots a graph of the conductivity of a soap solution as a function of the concentration one sees a kink also appear in this curve at the critical concentration. The surface tension, the specific gravity of the solution, the viscosity, all these properties show the same phenomena. (Fig. 2).

Weight normality
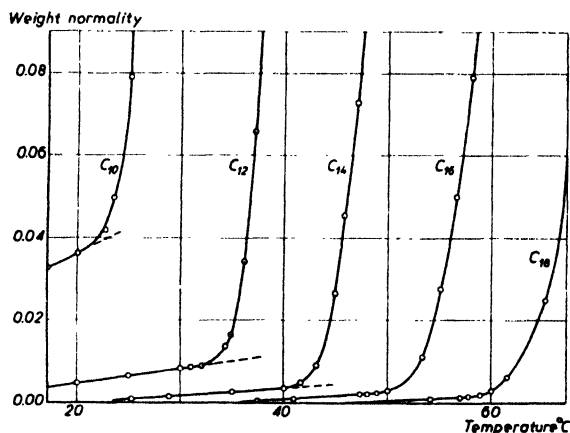
Fig. 3. Influence of the temperature on the solubility of Na-salts of higher alkyl sulphonates (according to TARTAR and WRIGHT, 1939).

TARTAR and WRIGHT [2] investigated the solubility of sodium salts of the higher alkylsulphonates at various temperatures. With rise in temperature the solubility slowly increases to rise rapidly at a particutar temperature. (Fig. 3).

Evidently the soap micelles formed are readily soluble; the temperature coefficient is extraordinarily large above the critical temperature.

The difference between the viscosity of soap solutions in water and alcohol (Fig. 4) demonstrates clearly that it is only in water that association to larger aggregates takes place [3].

This association can proceed so far that in certain cases elastic or gelatinous solutions are the result [4]. Thus the solutions of long carbon chains with aromatic sulphonate end groups are mobile with alkali ions, while copper and zinc ions give elastic solutions. The fact that the copper salt of cetylphenylether sulphonic acid is

[1] J. W. Mc BAIN in J. ALEXANDER, Colloid Chemistry, 1 (1926) 137. J. W. Mc BAIN, J. phys. Chem., 43 (1939) 671.

[2] H. V. TARTAR and K. A. WRIGHT, J. Amer. Chem. Soc., 61 (1939) 539.

[3] K. HESS, H. KIESSIG and W. PHILIPPOFF, Fette und Seifen, 48 (1941) 377.

[4] These systems exhibit two properties which lead to the name „elastic solutions". In the first place air bubbles which can be produced in the system by shaking, rise slowly but definitely upwards. With very small air bubbles this may take hours. From this we must therefore draw the conclusion that we are dealing with a liquid. If we rotate the test tube with the system through a small angle the air bubbles exhibit a vibratory motion. In other words the liquid has elastic properties (see H. G. BUNGENBERG DE JONG and G. W. H. M. VAN ALPHEN, Proc. Kon. Acad. Wetensch. Amst. 50 (1947) 849, 1011, 1227).

still just elastic at a temperature of 90° C in a 0.0002 % solution may serve as an example[1].

The colloidal particles, which are to be found in soap solutions, are formed spontaneously from the soap ions. This formation depends only on the concentration and on the temperature and not, as in hydrophobic colloids, on the previous history. The process is reversible and thus — as Mc Bain[2] has maintained for many years — the phase rule must be applicable to the soap colloids.

In the course of his very extensive studies of the soap boiling process many methods have been employed by him and his pupils; analysis of the phases, observation of the temperature at which phases appear or disappear, investigation with the microscope and the polarising microscope, determination of the vapour pressure and observations with the dilatometer. With the

Fig. 4. Viscosity of K laurate solutions of various concentrations in water and alcohol at 20° C (from Philippoff, 1941).

aid of these methods he manages to distinguish several phases in soap solutions, namely:

1°. Isotropic solution (nigre; lye). With these it must be noted that these solutions can be transformed into gels by addition of electrolytes, after which a separation into two layers can follow — the region of the coacervation of soap solutions studied later by Bungenberg de Jong and his pupils[3].

2°. Middle soap, an anisotropic (polarising microscope) viscous solution which contains still rather a lot of water (about 50%).

3°. Neat soap, also an anisotropic liquid with a lower water content (about

Fig. 5. Vapour pressure of Na-palmitate-water mixtures at 90° (Vold and Ferguson, 1938).

25%). Neither neat soap nor middle soap is miscible with the isotropic solution.

[1] G. S. Hartley, Kolloid-Z., 88 (1939) 22.
[2] J. W. Mc Bain, loc. cit., p. 684.
[3] H. G. Bungenberg de Jong, H. L. Booij and G. G. P. Saubert, Protoplasma, 28 (1937) 543, 29 (1938), 536, 30 (1938), 53; H. G. Bungenberg de Jong and G. G. P. Saubert, Protoplasma, 28 (1937) 498; H. G. Bungenberg de Jong, G. G. P. Saubert and H. L. Booij, Protoplasma, 29 (1938) 481, 30 (1938), 1.

4°. Curd fibres, soap crystals, which separate out of the solution at high concentration or on addition of much electrolyte.

The transition from the one phase into the other is — as was to be expected — characterized by the fact that various properties remain constant. See Fig. 5, for example, where the vapour pressure measurements of Vold and Ferguson[1] on Na-palmitate solutions are reproduced.

It may further be mentioned that Vold and Vold[2] distinguish six phases in the melting of Na-palmitate, in which four mesomorphic stages are to be found between the crystal and the isotropic liquid. They find the transitions between the phases at the following temperatures:

TRANSITION	TEMP.
curd fibres — sub waxy soap	117°
sub waxy soap — waxy soap	135°
waxy soap — sub neat soap	208°
sub neat soap — neat soap	253°
neat soap — isotropic liquid	292°

With the help of X-ray examination neat soap can be defined as a smectic liquid crystal, while sub neat soap is probably in an ordered smectic state.

However it is mainly the solutions of soap in water which interest us. When Krafft[3] drew the conclusion from his measurements of the rise of the boiling point of soaps that colloidal particles must be present in these solutions, he believed that the cause of this had to be sought in the hydrolysis of the soaps. Mc Bain[4] laid emphasis on the fact that hydrolysis only plays an important part in very dilute soap solutions, so that the reason for the appearance of colloidal particles has to be sought elsewhere. In the years 1913 to 1920 he worked on the comparison of

TABLE 1

EQUIVALENT CONDUCTIVITY AND HYDROLYSIS OF NA-PALMITATE SOLUTIONS AT 90 °C
(FREUNDLICH 1932[5], TABLE 50).

CONCENTRATION Na-palmitate in mol.	Λ	HYDROLYSIS in %
0.01	137.0	6.6
0.05	88.61	2.22
0.1	82.51	1.28
0.2	82.38	—
0.347	87.04	—
0.5	89.48	0.37
0.75	87.48	0.30
1.0	84.66	0.20
1.5	84.50	—

[1] R. D. Vold and R. H. Ferguson, J. Amer. Chem. Soc., 60 (1938) 2066.
[2] R. D. Vold and M. J. Vold, J. Amer. Chem. Soc., 61 (1939) 808.
[3] F. Krafft, Ber., 29 (1896) 1328.
[4] J. W. Mc Bain and H. E. Martin, J. Chem. Soc., 105 (1914) 957. J. W. Mc Bain and F. R. Bolam, J. Chem. Soc., 113 (1918) 825.
[5] H. Freundlich, Kapillarchemie, II (1932) p. 336.

investigations of the osmotic behaviour and the conductivity. We already know that the conductivity depends greatly on the concentration of the soap (see Table I). The figures for the hydrolysis already show that the latter plays no great part in the region of the large change in the conductivity.

Now the changes of the osmotic properties (lowering of the freezing point etc.) do not run parallel to the electrochemical properties and from these differences Mc BAIN [1] calculates the number of colloidal particles at various concentrations. Colloidal soap can exist in two forms: the anions can associate, but it is also possible that un-dissociated soap molecules form an aggregate. In a soap solution (for example Na-palmitate, NaP) the following particles can therefore occur:

in molecular solution   colloidal
    Na^+, P^-, NaP $(P^-)_n$, $(NaP)_x$

As far as the colloidal particles are concerned, there are three possibilities:
a.   The fatty acid anions are situated on the surface of a micelle of undissociated molecules.
b.   There are two independent micelles: an ionic micelle and a neutral colloid.
c.   The micelles mentioned in (a) and (b) occur together.

Mc BAIN [2] believes that all the properties of soap solutions can be explained in the following way. There are four regions in the concentrations of soaps which are characterized by definite changes in the composition. These four regions are the following:

1°.   In very dilute solutions the soap molecules are practically entirely dissociated; here the osmotic and the electrochemical properties will not deviate from each other.
2°.   After a critical concentration a fall of the conductivity occurs, with which a not proportionate decrease of the osmotic behaviour is associated. At this point then the dissociation must decrease and single sodium palmitate molecules will be produced. These molecules must be considered as the forerunners of the colloidal particle.
3°.   Then follows a region, in which the opposite takes place. Both the conductivity and the osmotic effect decrease — the second decrease is more rapid — and from this Mc BAIN concludes that neutral colloidal particles have been formed.
4°.   Finally the two properties rise again, but the osmotic effect remains far below the electro-chemical behaviour. These phe-nomena can best be explained by assuming the formation of strongly charged micelles — ionic micelles.

Fig. 6. Composition of Na-palmitate solutions according to Mc BAIN.

It is now possible to get some

[1] J. W. Mc BAIN, M. TAYLOR and M. E. LAING, *J. Chem. Soc.*, 121 (1922) 621.
[2] J. W. Mc BAIN, *J. phys. Chem.*, 43 (1939) 671.

idea of the amounts of the various particles from a comparison of the osmotic effect and the conductivity. Thus Mc BAIN[1] gave many graphs from which the composition of a soap solution at various concentrations can be read off (see for example Fig. 6).

The measurements of Mc BAIN on the salts of normal fatty acids could not be made below 0.01 n, because hydrolysis began to play an important part in very dilute soap solutions. This is shown very clearly in the measurements of EKWALL[2] who investigated the turbidity of soap solutions. At very low concentrations a region of turbidity occurs which can be attributed to products of hydrolysis. The turbidity disappears at higher concentrations after which a second region of turbidity makes its appearance.

Working with soaps, which contain sulphuric acid or sulphonic acid groups in place of carboxyl groups, furnishes the possibility of carrying out an investigation at much lower concentrations. There is almost no longer any question of hydrolysis with these strong acids, rather even the free fatty acids have a soap-like character. In fact the salts of these fatty acids exhibit the same properties as the ordinary soaps; here also a critical concentration can be assigned at which many properties change suddenly. But — as LOTTERMOSER and PÜSCHEL[3] found in measurements on salts of higher alkyl sulphonic acids — the critical concentrations, at which these sudden changes occur, lie much lower than the lowest concentrations, which Mc BAIN had measured. The conductivity of the acids studied by LOTTERMOSER and PÜSCHEL shows a considerable decrease at the following critical concentrations:

Alkali salt of C_{12} sulphonic acid . . . . 0.003—0.00625 n
,,   ,, ,, C_{14} ,,   ,, ... 0.0016—0.0025 n
,,   ,, ,, C_{16} ,,   ,, ... 0.0004—0.0006 n
,,   ,, ,, C_{18} ,,   ,, ... 0.0003   n

The temperature has only an insignificant influence on this critical concentration. Cations of the same valency also cause no shift of the point, but if one takes divalent in place of monovalent ions then the kink is displaced towards a lower critical temperature. The surface tension also begins to drop considerably in this region.

LOTTERMOSER and FROTSCHER[4] ask what is the origin of these large discrepancies with Mc BAIN's conception. The objection to Mc BAIN's ideas is that he works them out on the bases of the classical theory of dissociation and it is a great question whether this is permissible with the associated soap molecules and ions. LOTTERMOSER in fact sees the explanation of the sudden fall of the conductivity in the formation of neutral colloid, after which this neutral colloid slowly makes way for ionic micelles, which again results in a rise of the conductivity.

Is it not very improbable, asks HARTLEY[5], that two kinds of micelles, which exhibit exclusively electrical differences, occur in the same soap solution? In fact an attempt originated with this author to explain all the phenomena with the aid

[1] J. W. Mc BAIN, M. TAYLOR and M. E. LAING, loc. cit., p. 687. See also W. PAULI and E. VALKÓ, Elektrochemie der Kolloide (1929) p. 577 ff.
[2] P. EKWALL, Kolloid. Z., 77 (1936) 320.
[3] A. LOTTERMOSER and F. PÜSCHEL, Kolloid-Z., 63 (1933) 175.
[4] A. LOTTERMOSER and H. FROTSCHER, Kolloid Beih., 45 (1937) 303.
[5] G. S. HARTLEY, Kolloid-Z., 88 (1939) 22.

of one kind of micelle. It is very important that the solubilisation of organic substances insoluble in water begins at the critical concentration [1]. Micelles must therefore have already been formed at this point. This solvent action of soaps is too general to be explicable by means of the formation of complexes or crystalline compounds. Added to that the solubility of organic compounds in soap solutions is comparable with that in liquid paraffins. Is it any wonder that HARTLEY imagines the micelles formed as having the hydrocarbon chains of the soap ions entained in an irregular manner? A good argument for this idea is as follows: crystals of cetylpyridine chloride are heavier than water; a solution of the same salt in water is on the contrary lighter than water. This case reminds us of the melting of paraffin wax in which the specific gravity also decreases appreciably. That the salts of higher fatty acids show a similar tendency to form micelles is not due so much to the forces of attraction between the hydrocarbon chains as to the attraction of the water molecules. The long hydro-carbon chains are so to speak pushed out of the water. The longer the fatty acid, the greater the decrease of the free energy, and thus the lower the critical concentration for micelle formation will be. Equilibrium is established very rapidly in the micelle forma-tion; it is thus very improbable that compli-cated structures are formed.

The shape of the micelle can be deduced from simple considerations (HARTLEY 1939). With two associated fatty acid ions (Fig. 7a) the decrease of the boundary surface is rela-tively small. A spherical micelle in which the polar groups are directed outwards, is an energetically justified shape (Fig. 7b). This signifies that the micelles will have a particu-lar size. Indeed when the diameter of the micelle becomes much larger than twice the stretched chain of the fatty acid, polar groups must bury themselves in the non-polar part of the micelle and strong forces will resist this. Should the volume of the micelle be greater, then a flat shape would result (Fig. 7c). Against this HARTLEY has the objection that parallel hydrocarbon chains would be the result, which, according to him, is improbable. A second argument would be that the carboxyl groups repel each other with strong COULOMB forces.

Fig. 7. Possibilities of association of paraffin chain salts according to HARTLEY.

Now it must be possible to verify the theoretically deduced dimensions of the micelles with the aid of diffusion measurements. HARTLEY and RUNNICLES [2] measured the velocity of diffusion of cetylpyridine salts through a porous membrane and

[1] J. W. Mc BAIN, R. C. MERRILL Jr. and J. R. VINOGRAD, J. Amer. Chem. Soc., 63 (1941) 670.
[2] G. S. HARTLEY and D. F. RUNNICLES, Proc. Roy. Soc. A, 168 (1938) 420.

calculated the radius of the particles from this. If HARTLEY's idea is correct, the radius of the micelle must be independent of the concentration of the soap solution. Furthermore added salts must have only a slight influence on the dimensions of the micelles. This was confirmed by HARTLEY and RUNNICLES' investigations. Over a large range of concentration (from 0.002 to 0.05 n cetylpyridine chloride) and on addition of salt in various concentrations (from 0.04 to 1 n) the radius of the micelle appears to be fairly constant. The nature of the counter ion has a small but definitely measurable effect. The average values of the radius of the micelle in the following salts are:

acetate	$\nu = 24.2$ Å
oxalate	$\nu = 25.7$ Å
sulphate	$\nu = 26.6$ Å
chloride	$\nu = 27.2$ Å.

It is to be expected a priori that the size of the micelles will increase when organic compounds insoluble in water are added to the soap. These will indeed be taken up into the micelles. An attempt by HARTLEY and RUNNICLES to carry out these measurements failed through technical difficulties. The soap ions used by them are strongly absorbed on the negative glass of the apparatus — cetylpyridine chloride is in fact a cation soap — as a result of which the results of the experiments with added organic substances are greatly influenced by the degree of porosity of the membrane and are therefore valueless.

How then is the fall in the conductivity at the critical concentration to be explained? According to HARTLEY there are here two effects which operate against one another. The association of the fatty acid anions into a micelle would — when nothing else happened — undoubtedly lead to an increase of the conductivity (Mc BAIN effect). Indeed the single fatty acid ions have a fairly large mass in proportion to their charge and thus — in view of STOKES' law — a small electrochemical mobility. In the association of these anions into a spherical micelle the charge increases much more rapidly than the radius of the particle.

Thus the mobility must increase considerably to an extent dependent on the ratio of charge and radius. These considerations lead Mc BAIN[1] to attribute the fall of the conductivity to the formation of undissociated soap molecules, while LOTTER-MOSER[2] thinks that neutral colloid is produced in these circumstances. This conception by Mc BAIN is very improbable according to HARTLEY, since the considerable increase of the solubility of the soap above the critical concentration is then inexplicable (see Fig. 3). Also it is not clear to what the considerable solvent power for organic compounds could be attributed. Everything indicates that micelles are in fact produced at the critical concentration. Such a micelle of many fatty acid anions may however be considered more or less as a polyvalent ion and this makes it probable that deviations from the classical laws of ions occur. In the first place large DEBYE-HÜCKEL effects can be expected, while on the other hand a large part of the counter ions will be bound. This latter must be found by a comparison of the conductivity and the transport numbers (HARTLEY[3]). From such an investigation it

[1] J. W. Mc BAIN, *J. phys. Chem.*, 43 (1939) 671.
[2] A. LOTTERMOSER and H. FROTSCHER, *Kolloid-Beih.* 45 (1937) 303.
[3] G. S. HARTLEY, *Trans. Faraday Soc.*, 34 (1938) 1283.

appears that the conductivity of the paraffin chain ions (λ_p) increases considerably after the critical concentration. This is indeed one of the most direct pieces of evidence for the association of the fatty acid ions into a micelle. The conductivity of the counter ions (λ_g) decreases very considerably, because these are for a large part bound to the micelles and are transported to the similar electrode. Figure 8 demonstrates the results found very clearly. The fall of λ is to be imputed to the considerable fall of λ_g. It cannot be stated a priori which effect will win the conflict. In very strong fields the Mc BAIN effect can gain the victory as a result of which association will therefore be coupled with an increase of λ.

It is obvious that the binding of the counter ions to the micelles will be dependent both on the nature of the counter ions and on the nature of the polar group which is attached to the paraffin chain. The decrease of λ_g must indicate how strong the attachment of the counter ions to the micelle is. With cetylpyridine soap the binding of the monovalent ions appears to follow the rule:

$$I^- > Br^- > Cl^-$$

The binding of the I-ion is so strong that a yellow colour is produced. On the other hand

Fig. 8. Equivalent conductivity of cetyl pyridine bromide (λ), cetyl pyridine (λ_p) and bromide (λ_g). After the critical concentration λ_p increases (association), while λ_g exhibits negative values (the Br ions are carried along with the micelles). From HARTLEY 1939.

the mobility of the micelles is approximately the same in all cases, thus the size of the micelles must decrease in the same direction:

$$I^- > Br^- > Cl^-$$

Diffusion measurements (HARTLEY and RUNNICLES[1]) show up the like influences of the counter ions on the micelle size.

It has been known for a long time that the conductivity of soap solutions increases again at high concentrations. Here is the region where Mc BAIN[2] suspects the occurrence of ionic micelles. LOTTERMOSER[3] also expresses himself in favour of a similar opinion. HARTLEY[4] remarks that λ_p decreases (or does not rise rapidly) in the region, where λ_g rises and as a result of this the total conductivity increases. From this he arrives at the idea that the number of attached ions per micelle decreases. This would

---

[1] G. S. HARTLEY and D. F. RUNNICLES, loc. cit., p. 689.
[2] J. W. Mc BAIN, loc. cit., p. 690.
[3] A. LOTTERMOSER and H. FROTSCHER, loc. cit., p. 690.
[4] G. S. HARTLEY, Kolloid-Z., 88 (1939) 22.

imply that a retrograde dissociation occurs, at first sight a somewhat improbable hypothesis. HARTLEY[1] sensed the objections to this well enough and he slowly dropped the idea, so that he now thinks of the counter ions becoming more mobile. It is clear that there is a fairly large volume around the micelles where the concentration of counter ions is much greater than it is at a distance from the micelles. This must then have an influence on the resistance of the solution. Suppose that the concentration of an electrolyte in a tube is not uniformly distributed. In the one region the concentration $= C_0 (1 + A)$, in the other $C_0 (1 - A)$. The resistance will then be:

$$R = \frac{1}{kC_0 (1 + A)} + \frac{1}{kC_0 (1 - A)} + \frac{1}{kC_0 (1 + A)} + \cdots = \frac{e}{kC_0 (1 - A^2)}$$

If A becomes smaller, the conductivity will rise. According to HARTLEY something of the kind happens in a soap solution. At a low concentration the ratio of the ions in the clouds round the micelles to the ions in the solution is large; in other words, A is large and the conductivity small. If the concentration of the soap solution is increased, the above mentioned ratio and, with it, A decreases. The result will be a rising conductivity. Another anomaly of soap solutions also finds an explanation in this: a mixture of soap and electrolyte frequently gives a higher conductivity than that calculable from the sum of the conductivities of the components. An exaggerated picture well reproduces HARTLEY's views. Imagine that one has a mixture of spheres of a very well conducting material in a non-conducting liquid. The conductivity will naturally be zero. Now let us add a small amount of conducting medium and the result will be an abnormal rise in the conductivity. In the same way the soap micelles also have a large "internal" conductivity.

Recapitulating we can therefore say that HARTLEY believes that he can explain all the properties of soap solutions with one kind of micelle. This micelle as far as its structure is concerned stands more or less midway between the ionic micelle and Mc BAIN's neutral colloid[2].

§ 3. THE STRUCTURE OF THE SOAP MICELLES

HARTLEY's ideas can be applied very well to dilute soap solutions but in concentrated solutions other additional factors must be taken into account. HARTLEY's spherical micelle will not be able to explain the streaming double refraction of concentrated soap solutions. THIESSEN and TRIEBEL[3] investigated the streaming double refraction of Na-oleate at various concentrations and temperatures. This double refraction depends greatly on the concentration and on the temperature (see Fig. 9). At 50° the double refraction of a 15.3% solution of Na oleate has almost completely disappeared; while that of an 18% solution is still fairly high. Taken altogether it follows that anisotropic particles must be present in high concentrations of soaps.

---

[1] G. S. HARTLEY, Kolloid-Z., 88 (1939) 22.

[2] Later (J. W. Mc BAIN and A. P. BRADY, J. Amer. Chem. Soc., 65 (1943) 2072) Mc BAIN concedes that the term "neutral micelle" is not happily chosen. This micelle according to him will correspond to the lamellar aggregate indicated by HESS (see p. 694), which is much less strongly charged than the "ionic micelle". The "neutral micelle" is certainly not wholly uncharged.

[3] P. A. THIESSEN and E. TRIEBEL, Z. physik. Chem., A. 156 (1931) 309.

A negative double refraction was observed. This observation did not agree with what THIESSEN and TRIEBEL expected. They were of the opinion that rod-shaped particles occur in soap solutions. In that case a positive double refraction would be the result. It fits into their line of thought that the particles have a negative double refraction of their own (according to them the particles are crystalline). It must however be remembered that this investigation can just be evidence that the micelles are platelike (see the work of HESS, p. 694). In this latter case a negative double refraction would have to occur.

X-ray investigation should be able to un-ravel the crystalline structure of the soap micelles. THIESSEN and SPYCHALSKI [1] studied a 40% Na-myristate solution without result. No orientation in the particles could be per-ceived. If the solution is dried over sulphuric acid, they obtain the ordinary crystal structure of Na-myristate. Thus this investigation did not provide much information and they attempted another method. Rapid ultrafiltration ought according to them to separate the micelles without much change of shape from the equilibrium liquid, because — as their debatable hypothesis has it — change of shape in the micelles is a slow process. They then obtain a preparation which still contains only 20% water and a DEBYE-SCHERRER diagram of it shows the same picture as that of solid Na-myristate. From Mc BAIN's phase considerations it follows that this is rather logical; the equilibrium — contrary to the assertion of THIESSEN and SPYCHALSKI — is established rapidly and then this preparation settles into the "curd" phase. However, no conclusions may be drawn from this with regard to the micelles in less concentrated solutions. The structure of the micelle can be deduced from the X-ray diagrams of spun threads of hydrogels (Fig. 10). The hydrogels of soaps can be spun into threads, which are suitable for X-ray investigation after drying. Again they make the improbable hypothesis that nothing changes in this structure on drying. Investigation then shows that two

Fig. 9. Streaming double refraction of Na-oleate solutions in dependence on the temperature and concentration (THIESSEN and TRIEBEL, 1931).

Fig. 10. Structure of soap micelles according to THIESSEN and SPYCHALSKI, 1931.

constant spacings (7.5 and 4.9 Å) occur in all the soaps examined, while an inter-ference dependent on the length of the carbon chain appears as well. But once more it is improbable that these true crystals have anything to do with the soap micelles in dilute soap solutions.

[1] P. A. THIESSEN and R. SPYCHALSKI, *Z. physik. Chem.*, A. 156 (1931) 435.

There is nothing left but to investigate the X-ray diagrams of the soap solutions themselves. HESS and his collaborators [1] again devote their attention to the phenomenon of streaming double refraction and study the X-ray interferences of flowing oleate solutions. The DEBYE ring which corresponds to a spacing of 88 Å is then split into two arcs perpendicular to the direction of flow (see Fig. 11). The other interference (corresponding to 4.5 Å) remains unaltered. From a combination of these data one must draw the conclusion that the micelles are plate-like. In a stream they set themselves parallel to the direction of flow. That the short spacing does not alter signifies that the soap molecules must be perpendicular to the flat faces of the micelle.

Fig. 11. The DEBYE ring (88 Å) of a 16.9% Na-oleate solution (a) changes on flow into two sickles (b). According to HESS, KIESSIG, and PHILIPPOFF, 1938.

The X-ray interferences are sharper in soap solutions than in ordinary liquids. This makes it extremely interesting to compare the X-ray investigation with other properties of the soap solutions (HESS and coworkers [1 2 3]). When can X-ray interferences disappear with decreasing concentration of the soap?

1°. When micelles are no longer present in the soap solution.
2°. When the dimensions of the micelles become too small.
3°. When there is no longer any ordering in the micelles.
4°. When the concentration of the ordered phase is no longer large enough.

The most striking point of the X-ray measurements is that when an interference is found in the soap solutions which corresponds more or less to the length of two soap molecules, this spacing appears to be always longer than twice the calculated length of that soap molecule. In solid soaps a spacing is always found which is smaller than the double length of the molecules, because in this case the molecules are placed at a definite angle to the plane of the polar groups (Fig. 12). It appears very clearly from this that the micelles in an aqueous solution of soap have a different structure than the true soap crystals. From the fact that the spacing found is always longer than twice the length of the soap molecules — in some cases, for example Na-oleate, the spacing is much longer — we must conclude that water is built into the soap micelles. The following table gives an impression of the measurements of HESS [1] on many soaps:
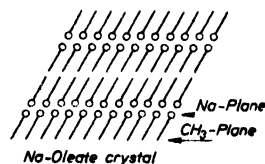
Fig. 12. Structure of the Na-oleate crystal.

[1] K. HESS, H. KIESSIG and W. PHILIPPOFF, *Naturwiss.*, 26 (1938) 184.
[2] K. HESS, W. PHILIPPOFF and H. KIESSIG, *Kolloid-Z.*, 88 (1939) 40.
[3] K. HESS, H. KIESSIG and W. PHILIPPOFF, *Fette und Seifen*, 48 (1941) 377.

TABLE 2

Soap	No. of C atoms	Conc. in % by weight	Long period in solution	Long period solid soap	Calculated length of 2 mols	Difference calc.-found	Difference in Å³	Molecules H₂O/ soap mol.
Na butyrate	4	31.5	—	—	13.0	—	—	—
,, valeriate	5	29.5	17.6	15.4	15.5	2.1	41.0	1.4
,, caproate	6	33.5	22.4	17.2	18.0	4.4	85.8	2.9
,, heptylate	7	29.4	24.6	19.2	20.6	4.0	78	2.6
,, caprylate	8	28.5	30.0	22.4	22.1	7.8	154	5.2
,, nonylate	9	26.5	31.6	23.0	25.7	5.9	115	3.9
,, caprate	10	27.4	37.7	26.8	28.2	9.5	185	6.2
,, undecylate	11	25.0	38.3	28.2	30.7	7.6	149	5.0
K laurate	12	33.4	43.5	29.5	33.3	10.2	199	6.7
Na oleate	18	18.7	78.0	44.3	48.5	29.5	575	19.4
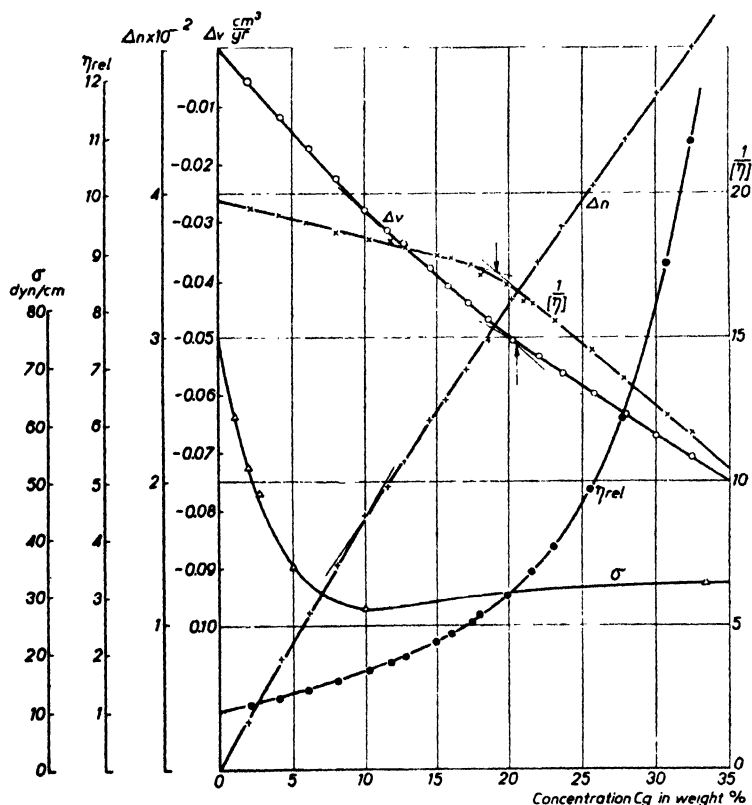
Fig. 13. Influence of the concentration on the viscosity (η_{rel}, ^{1}[η]), the specific volume (IV) the refractive index (Δn) and the surface tension (σ) of Na caproate solutions
(HESS, PHILIPPOFF, and KIESSIG, 1939).

Let us further illustrate the facts found for a few soaps. With Na caproate a considerable kink (at a concentration of 19% by weight) is observed in the viscosity curve (Fig. 13) (this kink is not clearly seen if η_{rel} is plotted against the concentration, but it is clearly visible when $1/[\eta]$ is used). Associated with this is a kink in the density curve at 20.5% while the X-ray interferences (corresponding to a spacing of 22.4 Å) disappear below 20 or 22%. The agreement between the numbers is excellent; a change in the state of solution takes place between 19 and 20.5% by weight of capro-ate, at which crystalline micelles are formed. It can also be perceived from Fig. 13 that a second kink occurs in some curves at a lower concentration (10% by weight). The surface tension for example has its minimum at this point. It seems obvious to think in this case of the formation of — unordered — HARTLEY micelles. The region below 10% by weight would then be the region of molecular solution.

Na-oleate is not so easy to align under this point of view. A clear kink again occurs in the $1/[\eta] - c$ curve and in fact at 7% by weight. X-ray interferences can still be found at lower concentrations (4—5% by weight). It must be pointed out that caproate and oleate show a remarkable difference in this respect. When we proceed to lower concentrations, the interferences of the short spacing (4.5 Å) disappear rapidly with the caproate while the others remain visible longer. This is exactly reversed with the oleate. Now we come to difficulties at a 4—5% by weight. The amount of the irradiated matter now becomes so small, that an X-ray interference will be very difficult to observe. Therefore the fact that the interferences are no longer visible at 4—5% by weight with the oleate, is no proof that ordering no longer

prevails in the micelles. The critical concentration for micelle formation naturally lies much lower. Where these micelles take on a crystalline character, on this point these experiments are therefore unable to give a decisive answer.

When — as appeared from the above mentioned investigations — water molecules are taken up in large amounts into the soap micelles, the question is interesting whether the amount of "bound" water depends on the concentration of the soap. This appeared from the investigations of HESS and his coworkers[1] to be in fact the case. If the concentration of the soap increases, the amount of water taken up falls and with it the long spacing determined by X-rays (Fig. 14)[2]. Since the soap
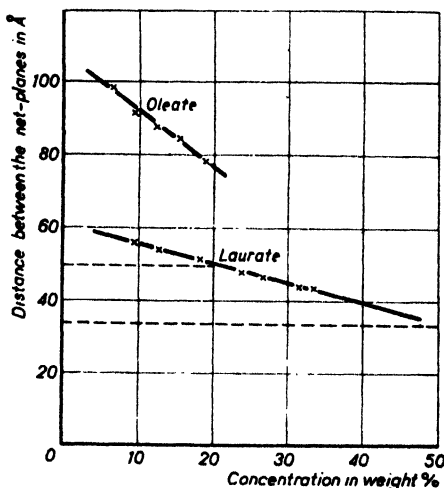
Fig. 14. Influence of the concentration on the long spacing in soap micelles (HESS, KIESSIG, and PHILIPPOFF, 1941).

molecules are arranged parallel, the water will be found between the polar groups

[1] K. HESS, H. KIESSIG and W. PHILIPPOFF, loc. cit., p. 694.
[2] This was confirmed by HARKINS et al. (W. D. HARKINS, R. W. MATTOON and M. L. CORRIN, J. Amer. Chem. Soc., 68 (1946) 220). They give the relation d = 21 — k log (1/c). See also S. Ross and J. W. Mc BAIN, J. Amer. Chem. Soc., 68 (1946) 296.

of the molecules (Fig. 15). An important question arises in this connection, namely what forces hold these thick layers of water between the soap molecules[1]. It is obvious that the long spacings in the micelles will be influenced by the added materials. Thus this spacing
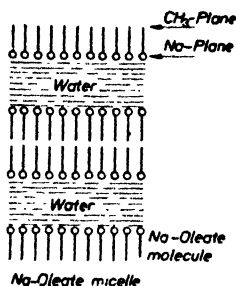
Fig. 15. Structure of the Na oleate micelle (HESS, KIESSIG, and PHILIPPOFF, 1941).
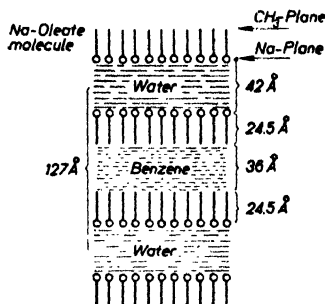
Fig. 16. Structure of a micelle in 9.12% Na oleate to which 0.79g of benzene has been added per gram of Na oleate (HESS, KIESSIG, and PHILIPPOFF, 1941).

increases steadily when benzene is added to a soap solution. The various facts agree best with the supposition that the benzene is taken up between the CH_3 groups of the soap molecules (Fig. 16). It is generally known that large amounts of meta-cresol can be taken up in soap solutions. Remarkably enough HESS finds that in this case the long spacing in the soap micelles does not increase but falls considerably. He cannot give an explanation (it is forgotten that meta-cresol will here function as an acid, addition of small amounts of mineral acids will undoubtedly also make the long spacing decrease).

HUGHES and his coworkers[2] point out that the electron density of the added compound influences the intensity of the interference lines. Hydrocarbons (with low electron density) intensify the interference lines; compounds with high electron density (for example bromobenzene) have the opposite effect.

There must be some kind of connection between the structure of the soap micelles and the mechanical properties of soap solutions. We have already seen that many soaps can form elastic or gelatinous systems and there is therefore every reason for establishing the connection between the properties and the X-ray observations. We have to thank PHILIPPOFF[3] in particular for attempts in this direction. To what extent — he asks himself — will the viscosity $[\eta]$ depend on the aggregation? Since the specific volume of the dissolved material will undergo little or no change through association of the particles, the shape factor must be influenced. In this connection various possibilities can be envisaged:

a. The aggregates are less asymmetrical than the original particles (lateral association). $[\eta]$ will then decrease.

b. The aggregates are more asymmetrical than the single particles (linear association). Naturally the value of $[\eta]$ rises in this case.

---

[1] See Chapter VI, Volume I.
[2] E. W. HUGHES, W. M. SAWYER and J. R. VINOGRAD, J. Chem. Phys., 13 (1945) 131
[3] W. PHILIPPOFF, Kolloid-Z., 96 (1941) 255; 100 (1942) 320.

c. The geometrical shape of the aggregates remains the same, in which case $[\eta]$ remains constant.

It is now important to investigate the course of $[\eta]$ at various concentrations of the soap solutions. In molecular soap solutions the value of $[\eta]$ lies between 0.04 and 0.06. This is not much more than the value of 0.025 which can be deduced for spherical particles. This is not at all astonishing as the hydrocarbon chain of a soap molecule will be greatly coiled up in aqueous medium. If the soap is in the micelle form, the values of $[\eta]$ become somewhat higher although they always still do not rise to really high values. We must therefore expect more or less spherical micelles and PHILIPPOFF gives the following representation of such a micelle (Fig. 17). Statistical packing together of these small micelles can, according to him, produce the X-ray interferences found by HESS.

On addition of electrolytes (for example KOH) other things happen however. KOH in high concentration added to laurate (PHILIPPOFF, 1942) causes a great rise in the viscosity. Now this viscosity is greatly dependent on the temperature. Not only this but a considerable structure viscosity also occurs, while the X-ray picture becomes obviously less sharp. PHILIPPOFF in fact is of the opinion that disordered aggregates are

Fig. 17.   Structure of the "Kleinmizellen" according to PHILIPPOFF, 1941.

produced by the addition of electrolytes. He then distinguishes three kinds of micelles:

1°.   Above the critical concentration there are produced micelles without structure viscosity, without streaming double refraction, independent of the temperature, giving no X-ray picture, having flat η_{rel}-c curves (HARTLEY micelles, "Kleinmizellen").

2°.   Then micelles which have all the characteristics of the first type but in addition show a clear X-ray picture (Groszmizellen).

3°.   Disordered aggregates, with a much higher viscosity, steep η_{rel}-c curve, dependent on the temperature, considerable structure viscosity and streaming double refraction but a poor X-ray picture.

STAUFF[1] opposes this representation. PHILIPPOFF's micelle model is not successful. It does not represent a minimum of the free energy, each successive micelle will place itself side by side with the previous one and so diminish the free energy while the minimum will only be reached with an infinite number of parallel coupled molecules. The words "Kleinmizelle" and "Groszmizelle" used by PHILIPPOFF (1941) are due to STAUFF[2], which is not to say that they both mean the same thing with these con-

[1] J. STAUFF, Kolloid-Z., 96 (1941) 244.
[2] J. STAUFF, Naturwiss., 27 (1939) 213; Kolloid-Z., 89 (1939) 224.

cepts. STAUFF expressly points out that the word crystalline must be handled cautiously in connection with soap micelles. He compares the X-ray photos of solutions of Na-tetradecylsulphonate at 20°ᶦ and 75° with that of liquid palmitic acid at 75°. In the first case three interferences (corresponding to 4.57, 3.98 and 38.3 Å) are observed. In the second case only two interferences occur (one constant which indicates a spacing of 4.6 Å, while there is in addition one which depends on the concentration of the soap). With palmitic acid also only one short spacing (4.6 Å) is to be found. The state of the hydrocarbon chains in a soap micelle must therefore correspond to that of a molten fatty acid. At 20° the Na-tetradecylsulphonate has crystallized out and therefore another interference makes its appearance, namely that of true crystals (Fig. 18). The criticism could be put forward that the disappearance of an interference band at 75° (b) is essentially the coalescence of the two bands of the crystal at 20° (a). The top of this band would then however have to lie

Fig. 18. Photometer curves of Na tetradecyl sulphonate (60%) at 20° (1) and at 75° (2), compared with liquid palmitic acid at 75° (3). d_1, d_2 and d_1' are the maxima of the interferences of the substances investigated, a and b are due to the NaCl used for calibration (STAUFF, 1939).

between the two tops of the original bands. It is clear that this is not so and thus that the state of the paraffin chains is quite different in the soap micelle from that in the crystal.

STAUFF and PHILIPPOFF are in agreement that the "Kleinmizellen" correspond to HARTLEY micelles. STAUFF also gives a model of these micelles which finds a place somewhat in between
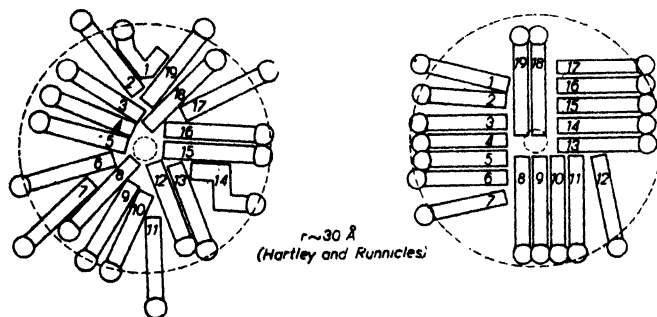
Fig. 19. Structure of the "Kleinmizellen" according to STAUFF, 1939.

the views of HARTLEY and of PHILIPPOFF (Fig. 19). The "Groszmizellen" which PHILIPPOFF looks upon as a statistical coalescence of "Kleinmizellen", are considered by STAUFF to be something quite different. According to him two possibilities present them-

selves with the "Kleinmizellen" (which we shall call HARTLEY micelles from now on):

1°. There are no attractive forces between these micelles (PHILIPPOFF [1]). Then the X-ray interferences must agree with the average distance between the micelles in the solution. This however does not come true. In 0.6 n Na-laurate the average distance is 65 Å, while 40 Å is found. Also this distance should decrease with $\sqrt[3]{C}$ which does not happen.

2°. There are attractive forces between the micelles.

Then there must be a minimum of energy, in other words, a limited "Groszmizelle" of a definite size is produced which is in equilibrium with the HARTLEY micelles. STAUFF (1941) [2] gives the following picture of this "Groszmizelle" (Fig. 20).

These "Groszmizellen" according to STAUFF are produced from associated HARTLEY micelles. A definite critical concentration cannot be detected; there is an equilibrium between the two kinds of micelle. The formation of the "Groszmizellen" begins at the minimum of the conductivity. STAUFF attributes the rise in the conductivity which then follows to the increase in the activity coefficient of the counter ions as the micelle concentration becomes smaller.
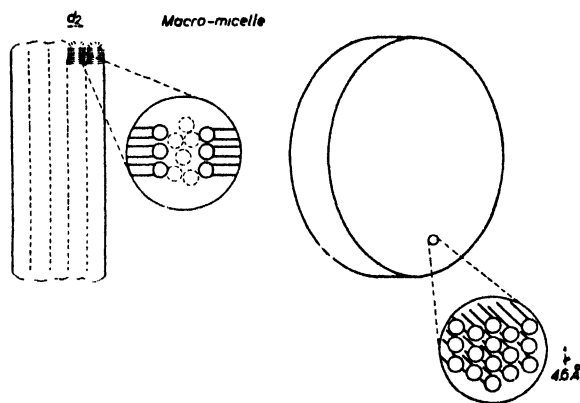
Fig. 20. Structure of the "Groszmizellen", a) cross section, b) the whole micelle (STAUFF, 1941).

For the present we shall restrict our criticism of STAUFF's representation to the statement that it is remarkable that his "Groszmizellen" — which have dimensions of approximately 250 Å · 2500 Å — cannot be seen in the ultramicroscope.

It may also be mentioned that oriented micelles occur in solutions of Na desoxycholate too [3]. However the orientation of the molecules is in this case not so easily established as in soap solutions. The lipids exhibit a similar orientation [4,5].

We may pause a moment at the question of what factors determine the detergent action of soap solutions. We will only examine in more detail the solubilisation, as a result of which the dirt is dissolved and taken up into the micelles (for a review of the most important factors in detergent action — solubilisation, emulsification, protective action, base exchange, suspending action — see an important article by Mc BAIN in *Advances of Colloid Science* [6]). HARTLEY [7] imagines that the organic material would be dissolved in the interior of the spherical soap micelles.

[1] W. PHILIPPOFF, *loc. cit.*, p. 697.
[2] J. STAUFF, *loc. cit.*, p. 698.
[3] J. W. Mc BAIN and A. P. BRADY, *J. Amer. Chem. Soc.*, 65 (1943) 2072.
[4] R. S. BEAR, K. J. PALMER and F. O. SCHMITT, *J. cell. comp. Physiol.*, 17 (1941) 355.
[5] K. J. PALMER and F. O. SCHMITT, *J. cell, comp. Physiol.*, 17 (1941) 385.
[6] J. W. Mc BAIN, *Adv. Colloid Sci.* I, (1942) 99.
[7] *loc. cit.*, p. 691.

However with the soaps with long hydrocarbon chains lamellar micelles will be present even at relatively low concentrations. A large amount of a compound insoluble in water can be taken up in these lamellar micelles — see the work of HESS[1].

It is probable that this is a very important factor for detergent action. It is to be expected that each factor which promotes the formation of lamellar micelles will through this increase the solubilisation at the same time. However it can be anticipated that the nature of the material to be dissolved will be very important; thus a compound with a strongly polar group will not be able to be admitted into the non-polar part of the micelles. In this case we expect an adsorption on the micelles or inclusion between the parallelly arranged soap molecules.

§ 4. COACERVATION OF SOAP SOLUTIONS

In the phase diagrams of soap solutions to which electrolyte has been added (see Fig. 21) there occurs a region in which the solution separates into two layers under the influence of the added electrolyte. This region has been profoundly studied by BUNGENBERG DE JONG and his coworkers[2]. BUNGENBERG DE JONG[3] has occupied himself these many years with the condensed phosphatide systems. The idea was that many biological problems (the rôle of calcium ions for the organism, ion antagonism, questions of permeability) could be studied in detail on these relatively simple models. The phosphatides play a great part in the protoplasmic membrane (OVERTON's lipoid
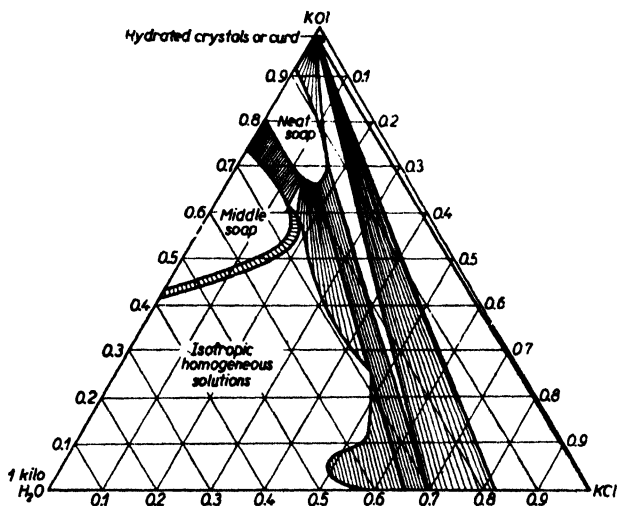
Fig. 21. Phase diagram of the system K-oleate, KCl and water at 20° C (Mc BAIN, 1926).

---

[1] loc. cit., p. 694.

[2] H. G. BUNGENBERG DE JONG, H. L. BOOIJ and G. G. P. SAUBERT, Protoplasma, 28 (1937) 543; 29 (1938) 536; 30 (1938) 53. H. G. BUNGENBERG DE JONG and G. G. P. SAUBERT, Protoplasma, 28 (1937) 498. H. G. BUNGENBERG DE JONG, G. G. P. SAUBERT and H. L. BOOIJ, Protoplasma, 29 (1938) 481; 30 (1938) 1. See also P. KOETS and H. G. BUNGENBERG DE JONG, Protoplasma, 30 (1938) 206 and S. ROSENTHAL, Diss., Leiden 1939.

[3] H. G. BUNGENBERG DE JONG, La coacervation; les coacervats et leur importance en biologie, Paris 1936.

theory) and BUNGENBERG DE JONG [1], WINKLER [2] and BOOIJ [3] suppose that this membrane
is a condensed phosphatide system. It is then logical to look for analogies between the
influence of various compounds on the permeability and that on the water content of
similar systems. We know that many organic non-electrolytes have a great influence
on the organism, in which TRAUBE's rule frequently comes to light. The influence
of these compounds on models was obviously a prominent part of the research.
Now phosphatides are however somewhat intractable. To obtain in this case a
separation into two layers (coacervation) — where the changes in the volume of the
colloid-rich layer and thus the water content are investigated — one must add to the
phosphatide sol a substance which increases the attraction between the non-polar
groups (for example triolein or cholesterol). It is certainly not impossible for a particular
organic non-electrolyte to have a quite different influence on the phosphatide system
in the absence of triolein or cholesterol from what it has when these substances have
been added. It is therefore desirable in the first place to study coacervates which are
produced without the addition of non-electrolytes. Soap coacervates then come
to mind for this purpose.

If one adds KCl to a solution of Na-oleate, the viscosity rises rapidly at a particular
concentration so that the system can even gelatinise. If more KCl is added a separation
into two layers begins, the upper one of which contains practically all the soap. With
rising KCl-concentration this coacervate layer becomes smaller and smaller.
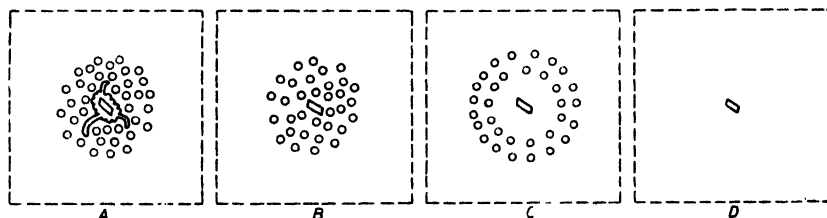
Fig. 22.   Vacuolization of an oleate coacervate under the influence of various substances.
(BUNGENBERG DE JONG and SAUBERT, 1937).

BUNGENBERG DE JONG and SAUBERT [4] investigated in the first place the influence
of many organic compounds on an isolated coacervate layer. A small crystal of the
substance to be investigated was put into a drop of coacervate and the reaction
could be followed under the microscope (contact method). This furnishes a clear
classification of organic non-electrolytes. The following figure (Fig. 22) shows the
various reactions at room temperature.

I A.   Around the small crystal there is produced a dense field of vacuoles
which expands continuously. In addition we see grains appearing — frequently

[1] H. G. BUNGENBERG DE JONG and J. BONNER, *Protoplasma*, 24 (1935) 198. H. G. BUNGENBERG
DE JONG and G. G. P. SAUBERT, *Protoplasma*, 28 (1937) 352.
[2] K. C. WINKLER, *Diss.*, Leiden 1938. K. C. WINKLER and H. G. BUNGENBERG DE JONG, *Arch.
Néerl. Physiol.* 25 (1940) 431, 467.
[3] H. L. BOOIJ, *Rec. Trav. Botan. Néerl.*, 37 (1940) 1.
[4] H. G. BUNGENBERG DE JONG and G. G. P. SAUBERT, *Protoplasma*, 28 (1937) 498.

birefringent — while finally myelin tubes [1] can make their appearance from them. A strong condensing action — the excess of solvent is separated in the form of vacuoles — must be the cause of these phenomena. Examples benzene, chlorobenzene and isoamyl urethane.

I B. Only a vacuole field is produced, thus in this case a weaker condensing action. Examples phenanthrene, p-dichlorobenzene and isobutyl urethane.

I C. First we see vacuoles produced around the crystal. This appears however to be a ring of vacuoles which slowly shifts. The crystal finally lies in a clear zone. In this case we must assume that low concentrations of the compound investigated have a condensing action, higher concentrations on the other hand an opening action — in the sense of favouring water-uptake. Examples propylurethane and hydroquinone.

I D. Microscopically there is nothing to see. This can however have many causes. The compound can have a weakly condensing to a strongly opening influence; in this experiment the results will always be alike.

When a preparation at room temperature still shows no changes after some time, we can demonstrate a differentiation into the various possibilities by raising the temperature. In this we make use of the fact that an oleate coacervate is condensed by rise of temperature, in other words it will vacuolise. The pictures, which are then observed, are collected in Fig. 23.

E   F   G   H
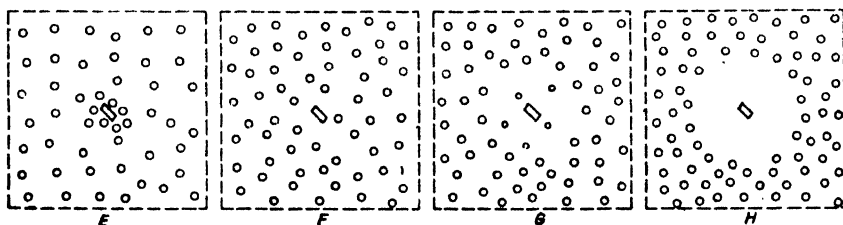
Fig. 23. If the preparation shows no vacuoles at room temperature (Fig. 22D), it is still possible for various images to appear on warming (BUNGENBERG DE JONG and SAUBERT, 1937).

II E. On warming a somewhat larger number of vacuoles are produced round the crystal than further out in the preparation. This indicates a weak condensing action, which was not observable at room temperature. On cooling the vacuoles around the crystal disappear the last, or they continue to exist. Example anthracene.

II F. No difference is observable — both on warming and on cooling — between the vacuoles close by and far away from the compound under investigation, in other words the substance has no influence. Examples chrysene, hexachlorobenzene.

II G. The vacuoles appear later in the neighbourhood of the crystal, while the density of the vacuoles is less there than further out in the preparation. A weak opening action is present in this case.

II H. No vacuoles are formed round the crystal, so that it comes to lie in a

---

[1] Birefringent tubes which grow from lipoid-containing material after the addition of water (for example, from myelin sheaths of nerves, from the protoplasma surface of many cells — by addition of alkali — and from isolated lecithin). The parallel orientation of the hydrocarbon chains causes the birefringence. See A. FREY-WYSSLING, *Submikroskopische Morphologie des Protoplasmas und seiner Derivate*, 1938, p. 62.

space as clear as glass. A strong swelling action must be the cause of this phenomenon. Examples: ethylurethane and acetanilide.

There is a disadvantage connected with this method. It is difficult to compare the actions of various compounds with each other because the solubility in the coacervate plays a great part. If a substance has no influence then this may certainly also be attributed to the insolubility of that substance. Thus the action of aromatic hydrocarbons decreases in a manner parallel to the increase of the boiling point. Does this in fact mean that the condensing action decreases in this direction? This seems improbable. The decrease — in the same direction — of the solubility is likely to be the explanation of the observed facts.

For the investigation of organic non-electrolytes which evaporate readily another method due to BUNGENBERG DE JONG and SAUBERT[1] is available. A drop of oleate coacervate is again to be found in a closed space and a small amount of organic substance (for example benzene, ethyl alcohol etc.) is then introduced into this space. The behaviour of the drop of coacervate is again observed microscopically. This method is not so suitable for a systematic investigation. Not only must the compounds to be investigated have a appreciable vapour pressure but also the solubility in the coacervate must not be too
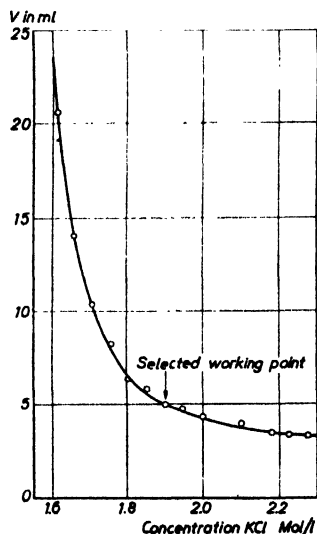
Fig. 24. Influence of KCl on the thickness of the layer of the separated oleate coacervate (BUNGENBERG DE JONG, BOOIJ, and SAUBERT, 1937).

small. Exactly as in the contact method the appearance or disappearance of vacuoles is in this case the characteristic by which the action of a particular compound is recognised.

The method of BUNGENBERG DE JONG, BOOIJ, and SAUBERT[2] is much more suitable for extensive
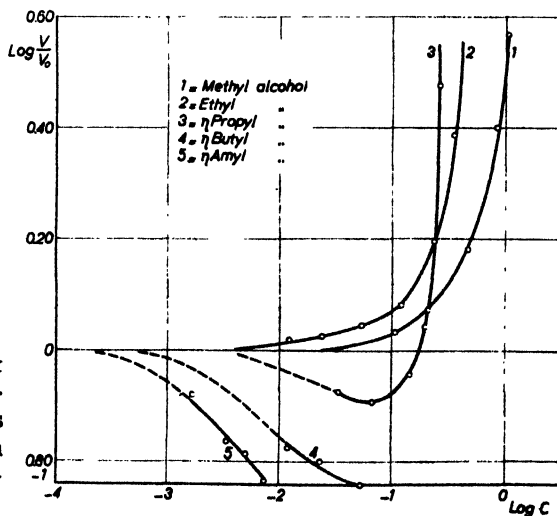
Fig. 25. Influence of the homologous series of the normal alcohols on the coacervate volume of the oleate (BUNGENBERG DE JONG, BOOIJ, and SAUBERT, 1938).

[1] H. G. BUNGENBERG DE JONG and G. G. P. SAUBERT, *Protoplasma*, 28 (1937) 329. See also p. 446 of this book.

[2] H. G. BUNGENBERG DE JONG, H. L. BOOIJ, and G. G. P. SAUBERT, *Protoplasma*, 28 (1937) 543.

systematic measurements. The principle of this procedure is as follows. When KCl is added to a Na-oleate solution the volume of the soap-rich layer[1], which separates out, becomes steadily smaller (Fig. 24). We choose a working point on the curve so obtained and then add — at constant KCl concentration — various organic compounds. Then we investigate the changes in the volume of the coacervate layer. Fig. 25 shows the influence of the first members of the homologous series of normal alcohols on the oleate coacervate. It is clear that the various actions already observable with the contact method have also been demonstrated here: methyl and ethyl alcohol have an opening action, butyl and amyl alcohol a condensing action while propyl alcohol assumes an intermediate place.

These alcohols have a comparable influence on lecithin coacervates which have a much smaller water content. It is only at high alcohol concentrations that complications appear which can be attributed to the fact that lecithin is an amphoteric substance.

A large number of organic compounds have been investigated as regards their influence on oleate coacervates and some general rules could be drawn up in connection with it.

1°. In homologous series of compounds with a polar group the concentration, at which an effect is attained, decreases with increasing number of carbon atoms.

2°. In such a series the action generally reverses at a particular chain length; the lower members have an opening action, the higher ones a condensing action.

3°. The influence which compounds with a hydrophilic group have on oleate coacervates can be decomposed into two opposite forces:
   a. a condensing action of the hydrocarbon chains,
   b. an opening action of the polar group.
   We gave (BUNGENBERG DE JONG, SAUBERT, and BOOIJ[2]) the following representation (Fig. 26).

Fig. 26. Representation of the course of events in the oleate coacervate. If the non-polar group is predominant, only condensation takes place (type I, for example amyl alcohol in Fig. 25). Afterwards a type can make its appearance (II) in which the same things occur at first as in the first type, followed by an opening action (propyl alcohol in Fig. 25). Finally there are substances which only exhibit an opening action (type III, methyl alcohol in Fig. 25.) (After BUNGENBERG DE JONG, SAUBERT, and BOOIJ, 1938).

Curve type *I*   Curve type *II*   Curve type *III*

Condensation   Swelling
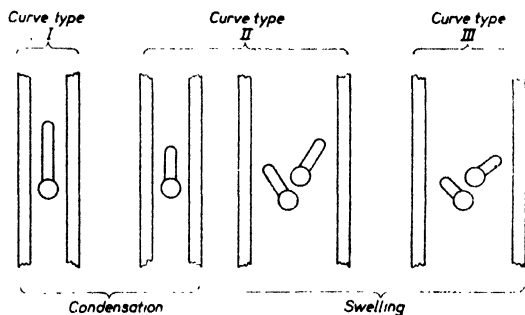
4*. The condensing action of the hydrocarbon chain increases with the length. With an equal number of carbon atoms the influence decreases with branching

---

[1] We use the word "soap-rich" only in a relative sense, for the absolute soap content of the "soap-rich layer" is always small. This follows directly from the phase diagram (Fig. 21). One can without difficulty make an oleate coacervate which contains only ½% of soap. This fact is very important for one's ideas on soap coacervates.

[2] H. G. BUNGENBERG DE JONG, G. G. P. SAUBERT, and H. L. BOOIJ, *Protoplasma*, 30 (1938) 1.

of the chain, with ring closure and on the transition from a saturated six membered ring to an aromatic ring.

5°. If a halogen is introduced into an aliphatic chain the result is a stronger condensation.

6°. Among the groups which favour an opening action the following can be included: the OH group, ether-, ketone- and ester-oxygen, etc.

7°. From the experiments it is possible to compare the action of polar groups with each other. Thus with the same number of carbon atoms, urethanes have a stronger action than primary alcohols and ketones a stronger action than secondary alcohols.

8°. In general the action of a compound is determined by the structure of the hydrocarbon chain (number, distribution and state of binding of the carbon atoms) and by the nature, the number and the position of the polar groups.

That these experiments can be of great value to biology was demonstrated by an investigation by BOOIJ[1] who in a comparison of the influence of many organic non-electrolytes on the germination of *Lathyrus* pollen found practically the same series which was already known for the action on oleate coacervates.

The above described method can only be used to investigate the influence of compounds which are fairly readily soluble in water. KOETS and BUNGENBERG DE JONG[2] describe another method by which insoluble compounds can also be investigated. They first dissolved these substances in oleic acid, after which they prepared oleate from this. By mixing this solution with blank oleate a series of different concentrations can be prepared. Aromatic and aliphatic hydrocarbons can also be investigated in this way. The rules given above for substances with a polar group no longer hold for exclusively non-polar compounds. Among other things the condensing action decreases in the direction hexane — heptane — octane. It is also remarkable that these compounds — although they do not contain a polar group — have an opening action at higher concentrations.
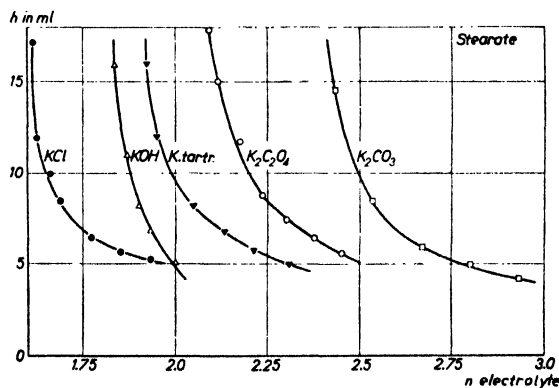
Fig. 27. The influence of various salts on the partial miscibility of stearate (ROSENTHAL, 1939).

The coacervate can then completely disappear as a result of which a very viscous system is produced. With the aromatic compounds the condensing action increases with the number of rings — thus in the direction benzene — naphthalene — anthracene. So it is remarkable that a strong condensing action is coupled with the phenanthrene structure as well as with cholesterol (in spite of the OH group) which again is reminiscent of biological analogies (hormones, steroids, carcinogenic hydrocarbons).

[1] H. L. BOOIJ, *Rec. Trav. Botan, Néerl.*, 37 (1940) 1.
[2] P. KOETS and H. G. BUNGENBERG DE JONG, *Protoplasma*, 30 (1938) 206.

The concentration, at which a coacervate is produced, depends greatly on the electrolyte used. Not only are the counter ions important but also the ions which carry the same charge as the fatty acid (Fig. 27). It is impossible to obtain an oleate coacervate at room temperature with sodium salts; floccules will be produced. The temperature exerts a great influence on the production of coacervates and in fact rise in temperature gives a reduction of the coacervate volume. In other words less KCl is necessary at high temperatures to obtain coacervation (Fig. 28). The oleate coacervates thus possess a negative temperature coefficient with respect to their solvent content. Also the influence of the organic non-electrolytes decreases relatively with rise in temperature.

A group of compounds with a polar group and a carbon chain — the fatty acids — have not yet been discussed. The effect of them is that they greatly decrease the volume of the coacervates (BUNGENBERG DE JONG, BOOIJ, and SAUBERT)[1]. However this action must be attributed exclusively to the acidic action; all acids investigated show precisely the same behaviour (formic acid, acetic acid, propionic acid, butyric
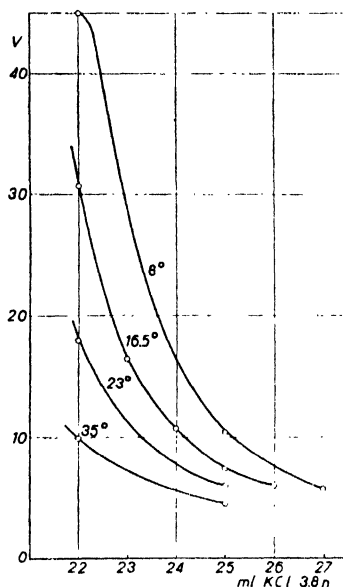
Fig. 28. The influence of the temperature on the coacervate volume of the oleate (BUNGENBERG DE JONG, SAUBERT, and BOOIJ, 1938).
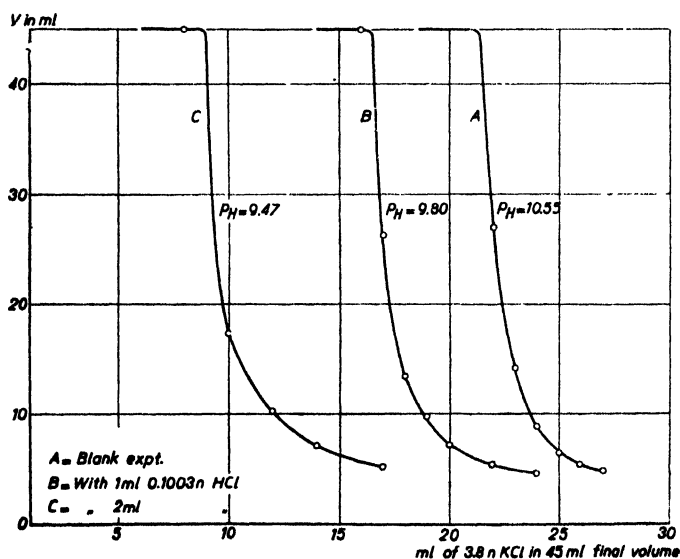
Fig. 29 The influence of the pH on the coacervation (BUNGENBERG DE JONG, BOOIJ, and SAUBERT, 1938)

acid, valeric acid and hydrochloric acid). The KCl curve is shifted very considerably

[1] H. G. BUNGENBERG DE JONG, H. L. BOOIJ, and G. G. P. SAUBERT, Protoplasma, 29 (1938) 536.

on the addition of a little acid (Fig. 29). The reason for this must be that the free fatty acid is formed, and this must have a strong condensing influence on the coacervate. This condensation can proceed so far that doubly refractive systems are produced.

ROSENTHAL [1] extended the researches of BUNGENBERG DE JONG to a large number of soaps. He compared stearate, palmitate, myristate, laurate, oleate, ricinolate and i-oxystearate. As the chain shortens more electrolyte is required to produce the partial miscibility of soaps. All the soaps were investigated at 60° C because no coacervation is possible at room temperature with many soaps. The following points may be noted with regard to his results. In the series of normal alcohols propyl alcohol formed the transition member between opening and condensing action when a oleate coacervate is the substrate. If, with the soaps, we go from 18 to 12 carbon atoms, ethyl alcohol slowly becomes the transition member. On introduction of hydrophilic groups in the fatty acid chain (oleate, i-oxystearate, ricinolate) this shift takes place in the same direction; propyl alcohol here finally exhibits a condensing action also.

An important question arises from these experiments on the coacervation of soap solutions: how can we connect up these experiments with the ideas on the micelles which must occur in soap solutions and how must we picture the structure of these remarkable soap-rich layers?

§ 5. THE CONNECTION BETWEEN THE STRUCTURE OF MOLECULE, MICELLE AND COACERVATE

With regard to many biologically active compounds — and in fact substances with important nonpolar molecules with or without polar groups — we can imagine that they are taken up in a layer of the organism containing phosphatide, for example the protoplasm membrane, and exert their action there. A particular group of this kind of compounds — the plant hormones, in particular the synthetic compounds — was investigated in detail by VELDSTRA [2].

These substances should be able, according to his working hypothesis, to increase the permeability of the young plant cell and thereby influence the growth. His hypothesis was now that the phytohormones ought to have an opening action on the protoplasmic membrane and in connection with this it was very well worth while to investigate the influence of these compounds on oleate coacervates. This action did in fact appear to be present. In this connection one thing must not be lost sight of; the free acids (for example naphthalene-acetic acid) naturally have a condensing action like all acids. But these experiments must be carried out at high pH, at which the phytohormones are completely present as ions.

This opening action of certain organic anions makes it extraordinarily interesting to investigate a group of compounds which had not yet been examined — namely the anions of the homologous series of normal fatty acids — for their influence on soap coacervates. The method was as follows: the fatty acid to be investigated was dissolved in an oleate solution to which a small excess of alkali had been added. Mixtures of this with a blank oleate solution could then be made of all concentrations of the fatty acid anion. The mixtures of oleate and fatty acid anion were brought into the coacervate region in the usual way with the aid of KCl: In so doing it is

[1] S. ROSENTHAL, Diss., Leiden 1939.
[2] H. VELDSTRA, Enzymologia, 11 (1944) 97, 137.

brought to light that the required amount of KCl is dependent on the amount of fatty acid anion present (Fig. 30). The shift which the KCl curve undergoes — it is obvious that a shift to-
wards higher concentrations corresponds to an opening action — is almost propor-tional to the amount of fatty acid anion. It is how-ever noteworthy that some fatty acid anions have a strong opening action (for example undecylic acid) while others (for example palmitic acid) have very little influence. The surpri-sing result of this research (BOOIJ and BUNGENBERG DE JONG [1]) is set out in Fig. 31.
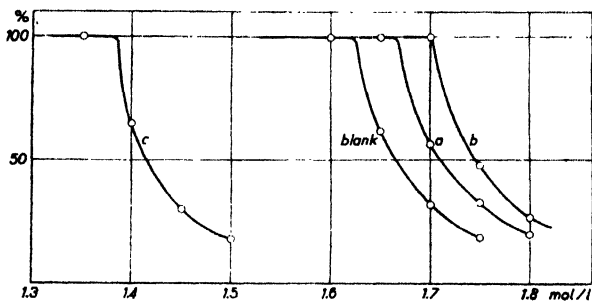
Fig. 30. Addition of palmitate (a, b) to an oleate coacervate causes the amount of KCl required for coacervation to rise; compare the action of a condensing substance, benzyl alcohol, c (BOOIJ and BUNGENBERG DE JONG, 1948).

The action of the lower fatty acid anions is practically zero at low concentrations. An opening action just begins at the caproate and rises considerably up to the undecylate. After this the action again falls to rise considerably with the higher fatty acid anions.

The kink which the curve exhibits and which has its trough at 15 carbon atom.
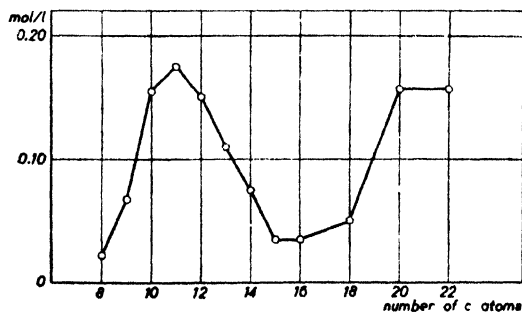
Fig. 31. To investigate the influence of the homologous series of the nor-mal fatty acids on an oleate coacervate the displacement of the KCl line, caused by 0.5 m.mol of added substance, is plotted against the number of carbon atoms of the added fatty acids. One works in a basic medium and thus the influence must be attributed to the fatty acid anions (BOOIJ and BUNGENBERG DE JONG, 1948).

is not so very remarkable. Indeed addition of a little oleate to an oleate coacervate has not much influence on the shift of the KCl curve. A fatty acid anion with a different number of carbon atoms does have an influence. Now it is also a fact that the smaller fatty acid anions are so readily soluble that they are practically not taken up in the soap micelles. The first part of the curve — from 6 to 11 carbon atoms — is thus to be attributed to the steadily rising amount of fatty acid anions which are taken up in the soap micelles in this direction. The idea which comes to the fore from a consideration of this curve, is that a certain structure prevails in the oleate coacervate and that anions with a longer or shorter paraffin chain

---

[1] H. L. BOOIJ and H. G. BUNGENBERG DE JONG, Bioch. et Bioph. Acta, in press.

do not fit into this structure, as a result of which this structure is distorted.

If this idea is correct then a similar curve ought to appear with a stearate coacervate in which the minimum should lie at 18 carbon atoms. This prediction could be confirmed experimentally (Fig. 32). That the minimum does not lie at 18 carbon atoms with the oleate coacervate, must naturally be attributed to the cis-configuration of the paraffin chain as a result of which the latter does not "pull its full weight". It is to be anticipated that the minimum will in fact lie at 18 carbon atoms with the elaidinate.
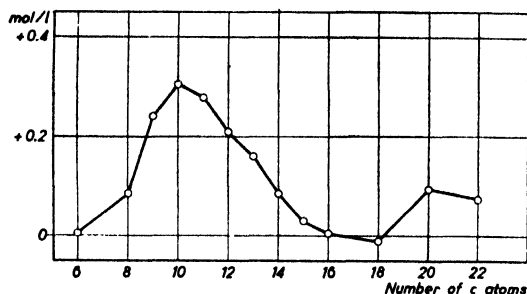
Fig. 32. The influence of the homologous series of fatty acid anions on a stearate coacervate. (see Fig. 31).

In an investigation of the influence of fatty acid anions on the temperature of partial miscibility of a Na-desoxycholate coacervate, the influence is also seen to begin at the caproate. After this a sharp rise of the action to a constant level follows (Fig. 33). There is no question in this case of a pronounced minimum in the curve such as was found with oleate and stearate coacervates. Perhaps there is a small minimum at about 14 carbon atoms. This would then indicate a certain degree of order in the desoxycholate micelles[1], which can be disturbed by fatty acid anions, but on which the myristate ion acts relatively less strongly than the other long fatty acid anions. This then must mean that the desoxycholate molecules also are arranged parallel in the micelles and that the thickness of a single layer of these molecules will be approximately as large as the length of a myristate molecule. This idea of an order in the soap coacervate fits very well with the data furnished by X-ray investigation. In fact HESS and coworkers[2] and STAUFF[3] also arrived at the conviction that there are ordered systems present in soap solutions in which the soap ions are aligned more or less parallel. There is, however, one objection, namely the idea of PHILIPPOFF[4] that unordered aggregates are produced by addition of KOH to soap solutions.

Fig. 33. The influence of the homologous series of fatty acid anions on a desoxycholate coacervate (see Fig. 31).

The question therefore arises whether we may assume ordered systems in the

[1] This was also made plausible in another way, see p. 700.
[2] K. HESS, W. PHILIPPOFF and H. KIESSIG Kolloid-Z., 88 (1939) 40.
[3] J. STAUFF, Kolloid-Z., 89 (1939) 224.
[4] W. PHILIPPOFF, Kolloid-Z., 96 (1941) 255.

soap coacervates, where in general still more electrolyte is necessary for their pro-
duction. The answer to this question must be sought with the aid of a combination
of the data as yet known on soap solutions.

As a starting point we choose the observation[1] that with the caproate the
X-ray interference corresponding to the short spacing (4.6 Å) disappears first
with decreasing concentration, while with the oleate this interference is just the
one which is maintained the longest. Apparently — according to HESS et al.,
1939 — the caproate micelle is rapidly broken up in the sideways direction on reducing
the concentration; they will be rods in stead of sheets. In the valeriate no 4.6 Å
spacing is found even in concentrated solutions, but the long spacing of 17.6 Å is found.
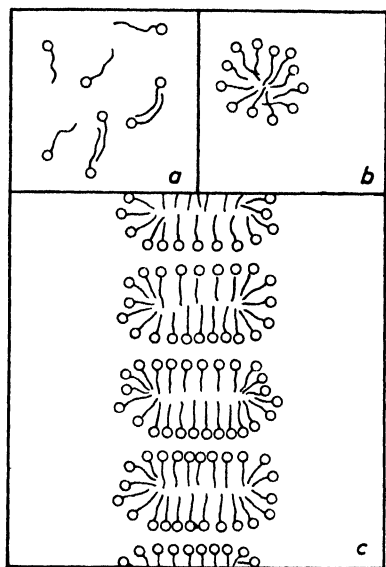
Fig. 34. Occurrence of micelles in strong-
ly concentrated solutions of Na valeriate.
Only the X-ray interference correspon-
ding to the long spacing appears, as the
micelles do not grow big enough for the
short spacing to appear.
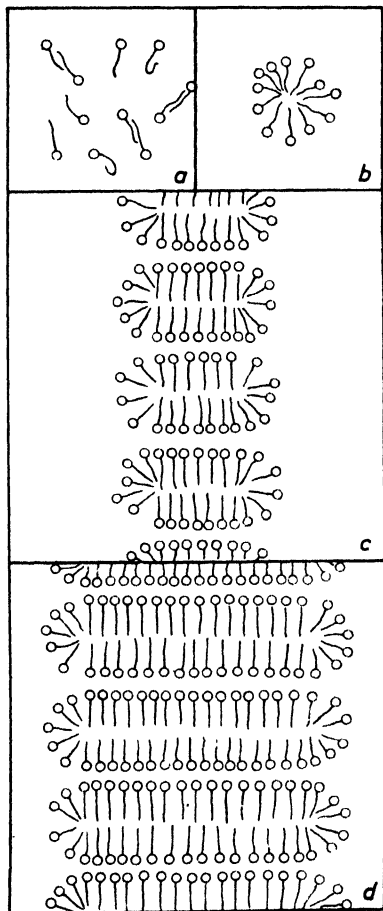
Fig. 35. With solutions of Na caproate the short
spacing will eventually also be observable, in other
words, flattened micelles are produced.

This must mean that there is no question or hardly any question of a parallel alignment
of the paraffin chains. If we attempt to make a picture from these discoveries we
arrive at the following idea of the structure of these soap solutions (Fig. 34 and 35).

[1] HESS et al., Kolloid Z., 88 (1939) 40.

At the critical concentration of the valeriate (see Fig. 34, transition from a to b) the HARTLEY micelles are formed, and these perhaps become somewhat flattened with increasing concentration and chains will start to form (Fig. 34c) whereby the X-ray interference of a long spacing makes its appearance.

The caproate behaves at low concentrations in a manner exactly comparable to the valerate (Fig. 35 a, b, c). Obviously in this case the flattening of the micelle at the high concentrations goes somewhat further (Fig. 35d), as a result of which a short spacing, namely the cross section of the paraffin chains (4.6 Å) becomes visible.

There are two forces which control the equilibrium of the soap micelles in aqueous solution. The negative polar groups repel each other, while the paraffin chains are drawn towards each other by LONDON-VAN DER WAALS forces. Undoubtedly these latter forces are not strong but it must be borne in mind that the free energy decreases very considerably if the paraffin chains are not in contact with the water. This decrease of the free energy is the most important contribution to association. We can investigate the possibilities with the aid of a schematic picture of two associated soap ions (Fig. 36).

At low concentrations — above the critical concentration however — the HARTLEY micelle is undoubtedly the most probable form. Indeed the dissocation is large in this case, force A therefore correspondingly large and the result will be a spherical micelle. If now the concentration becomes higher, the dissociation slowly diminishes, force A decreases, which will have the result that the size of the micelle increases. In the process the polar groups cannot dive into the micelle and so a flattened micelle must be produced. Apparently attractive forces must exist between these micelles since certain X-ray interferences are found which are rather larger than twice the length of the stretched chain of the molecules. This spacing — as we know from the work of HESS et al.[1] — depends on the concentration of the soap and the amount of water between any two micelles and depends on the length of the paraffin chain.

Fig. 36. Two opposing forces determine the structure of the soap micelles: A) the repulsive Coulomb forces of the charged groups and B) the attractive forces (LONDON-V. D. WAALS forces) between the carbon chains.

Apparently therefore force B (Fig. 36) has something to do with the number of water molecules to be found between the polar groups. This appears somewhat peculiar and it is still stranger that this number of water molecules is clearly influenced by the fact whether the carbon chain contains an even or an odd number of carbon atoms (see Table p. 695). Here we are reminded of the melting points of the fatty acids where the same fact plays an important part. We also see it occur in the experiments of DONNAN and POTTS[2] on the interfacial tension of soaps (in this connection see the discussion given by FREUNDLICH[3]).

The back-ground of this remarkable phenomenon will in fact have to be sought in force B. Naturally this force is stronger in fatty acids with an even number of carbon atoms (higher melting point) than in fatty acids with an odd number. The tendency to the parallel alignment of the chains — and thus to flattening of the

---

[1] K. HESS, H. KIESSIG, and W. PHILIPPOFF, Fette und Seifen, 48 (1941) 377.
[2] F. G. DONNAN and H. E. POTTS, Kolloid-Z., 6 (1910) 208.
[3] H. FREUNDLICH, Kapillarchemie, II (1932) p. 485.

micelle — depends directly on *B*. The consequence of this flattening is that the micelles carry a high negative charge on the flat sides — that is to say the charge per unit surface area is high. With the same dissociation the distance between two micelles will then rise on increase of force *B*.

This force will again increase on increase of the number of carbon atoms or on transition from a member with an odd number to one with an even number of carbon atoms (Fig. 37). Naturally with decreasing dissociation — increasing concentration — the distance becomes steadily smaller.

When the dissociation decreases considerably this has still another effect as a consequence. The force *A* (Fig. 36) continues to decrease further, the tendency of the micelles to round

Fig. 37. With a caproate micelle the layer of water between the polar groups will be much less thick than with an oleate micelle because the oleate micelle is much more flattened.

themselves off becomes steadily smaller. Eventually — with the inclusion of more soap ions — plate shaped structures are produced, which exhibit the well known phenomena of streaming double refraction in high concentrations of oleate solutions (THIESSEN [1]; HESS [2]).

When we start from the HARTLEY micelles there are thus two possibilities of association: a loose association of the micelles and an increase in the size of the micelles, whereby plate-shaped aggregates are formed. It must be possible to understand all the phenomena of soap solutions with the aid of these two types of association.

When we add KOH to a soap solution — of fairly high concentration — the viscosity falls and after a minimum rises very steeply (see Fig. 38). The X-ray interference of the long spacing remains the same, thus the thickness of the micelles does not increase or decrease [3]. The intensity decreases however, which in our opinion indicates that in the first place the association of the micelles decreases [4].

Fig. 38. Influence of KOH on the viscosity (A) and on the X-ray interference (B) of a 18.1 % K-laurate solution (HESS, KIESSIG, and PHILIPPOFF, 1941).

[1] P. A. THIESSEN and E. TRIEBEL, *Z. physik. Chem. A*, 156 (1931) 309.
[2] K. HESS, H. KIESSIG, and W. PHILIPPOFF, *Naturwiss.*, 26 (1938) 184.
[3] K. HESS, H. KIESSIG, and W. PHILIPPOFF, *Fette und Seifen*, 48 (1941) 377.
[4] W. D. HARKINS, R. W. MATTOON, and M. L. CORRIN, *J. Amer. Chem. Soc.*, 68 (1946) 220, similarly find a decrease of the intensity of the long spacing when they add KCl to a laurate solution. Their observation that addition of NaCl actually results in a stronger intensity of the interferences is apparently in conflict with our ideas. However true Na-laurate crystals are then produced (they observe clouds or a small amount of precipitate, whereby, in our opinion, the greater intensity is explained).

One would not expect this a priori. Indeed the dissociation of the polar groups decreases, thus the supposed decrease of the association of the micelles cannot be attributed to this. A consideration of a very schematic picture of the course of the events furnishes the solution of the poblem. We know (Fig. 39a) that an association of the soap micelles occurs in a soap solution of a fairly high concentration. On the other hand it is probable that large plate-shaped micelles are produced (Fig. 39c)
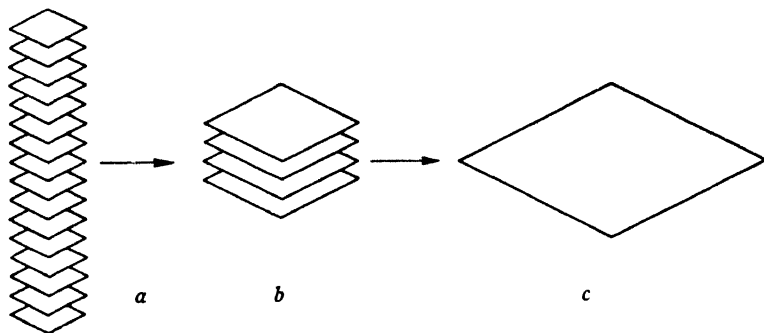
Fig. 39. Influence of KOH on the micelle shape of a concentrated soap solution.

after the addition of a considerable amount of electrolyte. At the transition from the one state to the other a phase must be passed in which the particles are less asymmetrical (Fig. 39b). It is clear that the viscosity will exhibit a minimum there.

The primary fall in the viscosity is relatively small in comparison with the rise which larger amounts of electrolyte cause. Since the formation of the large plate-shaped micelles must be the result of the reduction in the dissociation it cannot be otherwise than that the nature of the added electrolyte will be important. It would be worth while investigating the influence of various ions on the mobility[1].

Naturally the micelles formed will not be as flat as shown schematically in Fig. 39c. In spots with a large dissociation the surface will ꝺe curved (see Fig. 36; large force A). The undissociated spots form the places where the micelles can be attached to each other. As a result a sort of network is formed and the solution acquires its elastic and strongly viscous properties. It cannot be too strongly emphasized that we are dealing here with a dynamic system. The smallest air bubble rises slowly but surely upwards in such a soap gel. This means that the mutual attachment of the micelles is of relatively short duration.

At still higher electrolyte concentration the number of points of attachment on the micelles — which we can in this case consider as a sort of macromolecules — becomes continuously greater. The consequence of this is that the micelles will also adhere internally (to use a metaphor: a sheet of paper becomes a crumpled ball) and the viscosity again decreases.

In the long run the ball of "macromolecules" (= micelles) will contract so far

---

[1] The work of Mc BAIN has already been mentioned — p. 700 — from which it was concluded that the solubilisation of organic compounds is probably mainly due to the lamellar micelles. The fact that addition of electrolytes raises the solubilisation can be regarded as support for our hypothesis that electrolytes favour the formation of lamellar micelles.

that it no longer completely occupies the volume of the liquid. A coacervate separates, the volume of which decreases steadily with increasing concentration of electrolyte. Fig. 40 illustrates this idea. In dilute soap solutions the transition from A to B will not be found because the state A does not exist there.
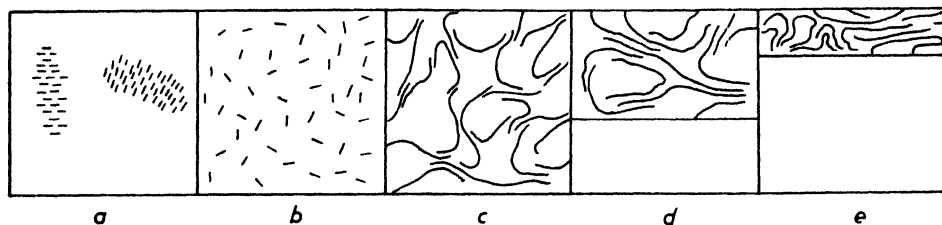
Fig. 40.   Micelle shapes in coacervation.

To elucidate this it is good to look again at the phase diagram of a soap solution (Fig. 21). There are many possibilities in the investigation of soap solutions (see Fig. 41). Many investigators study the changes occurring when the soap concentration is changed and this means that the lines A or A' are followed. In the investigation of coacervation the soap concentration is kept constant, in other words, one moves in the direction B. We have already depicted in Fig. 40 how we picture the changes in this direction.

Now something more about the structure of the micelles in soap solutions not containing electrolyte on alteration of the concentration (direction A). With a soap with a somewhat long paraffin chain (for example oleate) the formation of flat micelles takes place rather rapidly (Fig. 42a). The dimensions of these micelles will rise at higher concentrations through a decrease of the dissociation, but before this is an important factor another event first takes place. The micelles will align themselves with respect to each other, first still in small groups (b), later throughout the whole soap solution.

Fig. 41.   Possibilities in the study of soap solutions.

A solution of aligned (relatively small) micelles is produced; this solution will be birefringent (middle soap). If we look upon the soap micelles as the units in this system, then we can here speak of a nematic state (Fig. 42c). Then the decrease of the dissociation results slowly in an enlargement of the micelles, from which neat soap finally results. The state in neat soap is thus somewhat reminiscent of the state in a strongly condensed coacervate. A consideration of the tie lines in the coacervate region of the phase diagram furnishes a pretty indication in this direction (see for example the diagram for laurate in an article by Mc BAIN, BROCK, VOLD and VOLD[1]). It is possible to make out from the phase diagram how much water is really bound by the soap micelles and how much water is immobilized

---

[1] G. W. Mc BAIN, G. C. BROCK, R. D. VOLD and M. J. VOLD, *J. Amer. Chem. Soc.*, 60 (1938) 1870.

by the network (this is therefore really "sponge water"). This is done by studying the direction of the tie lines in a region of partial miscibility (see this book, BUNGENBERG

a   b   c   d   e
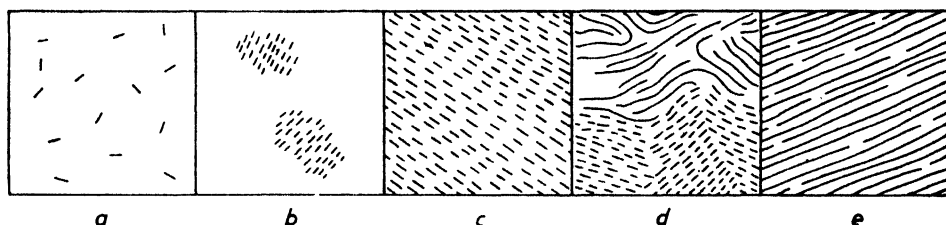
Fig. 42.   Changes in the micelle shapes on increase of the soap concentration.

DE JONG, Chapter VIII). If we do that in the case of soap coacervation we see that the tie lines point towards the region of neat soap. A clear indication that there is agreement between the micelles of neat soap and of the coacervate.

We cannot give a complete picture of the possibilities of the theory developed here. We may add a few examples by way of demonstration:

1°. Addition of acid to a soap solution causes the KCl concentration required for the coacervation to fall sharply (Fig. 29). This is a logical consequence of the fact that the number of uncharged spots on the micelle — and therewith the extension in the plane of the particle — increases. The large and flat micelles required for coacervation are formed more rapidly.

2°. The viscosity of many soap solutions exhibits anomalies which point to structure viscosity on addition of electrolytes (PHILIPPOFF[1]). It is obvious that nothing else could be expected with the picture we have given (Fig. 39).

3°. At room temperature no coacervates can be made with many soaps (stearate, palmitate etc.) for the simple reason that these soaps crystallise. This must be attributable to the magnitude of force B (Fig. 36). When now it is possible for us to "unloosen" the paraffin chains, coacervation will in fact be able to occur. Thus BUNGENBERG DE JONG (unpublished communication) was able to make a turbid, crystallised Na-stearate solution perfectly clear by adding menthol. We must imagine that the crystals are slowly transformed into micelles. Afterwards coacervation could be obtained with the aid of KCl.

4°. If one adds an electrolyte to a soap solution of lower concentration one sees no lowering of the viscosity such as happens when one starts from a higher soap concentration (BUNGENBERG DE JONG and VAN ALPHEN[2]). This is neither to be expected since at lower concentration the soap micelles are hardly associated (Fig. 42a). With the addition of electrolyte the chance of a loose association of the micelles becomes greater. Then we are in for the same thing as already demonstrated in Fig. 39. At a given moment an anomaly occurs which must here also be attributed to a transition from a loose association of the micelles into the large flat micelles (Fig. 43.) From this point onwards the soap solution can also be spun into threads.

[1] W. PHILIPPOFF, Kolloid-Z., 96 (1941) 255.
[2] H. G. BUNGENBERG DE JONG and G. W. H. M. VAN ALPHEN, Proc. Kon. Acad. Wetensch. Amst. 50 (1947) 849, 1011, 1227.

5°. In general the addition of an organic non-electrolyte can act in two ways. A soap coacervate can be condensed, in which case an opening action usually follows at higher concentrations. That this is not a consequence of the polar group, which is "smuggled" by the paraffin chain of, for example, amyl alcohol into the soap micelle, is to be seen already from the fact that octane, heptane and similar substances have the same influence (KOETS and BUNGENBERG DE JONG)[1]. The old ideas on the opening and condensing action (see Fig. 26) must in fact be abandoned. Perhaps in the first place — by addition of a small amount of an organic non-electrolyte — the tendency to the formation of larger flat micelles can be favoured because the micelles stick to each other sideways. At a given moment another tendency will begin to predominate: the effects, dealt with in the two previous paragraphs, which lead to a loose association of micelles, occur[2]. This
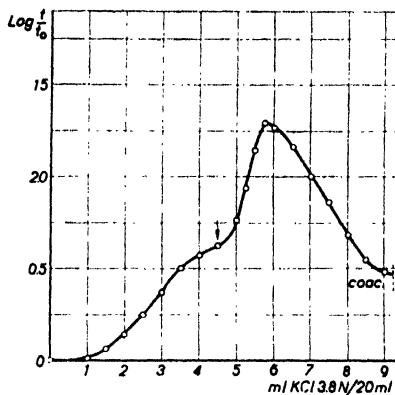
Fig. 43. Influence of electrolyte on the viscosity of a soap solution of low concentration. (BUNGENBERG DE JONG and VAN ALPHEN, 1947).
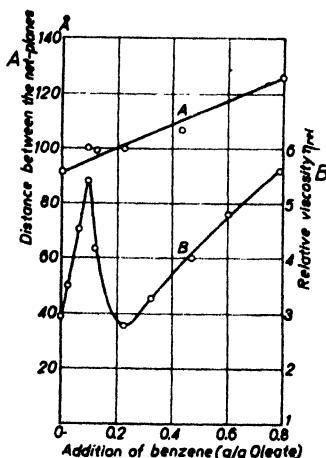
in fact is made manifest very clearly in the viscosity of such a system (see Fig. 44). While the thickness of the micelles gradually increases (benzene is thus taken up in the micelles between the CH_3 groups), the viscosity undergoes considerable changes on the addition of benzene[3]. A sharp rise — corresponding to the condensation of the coacervate — is first observed. The subsequent drop must then be the consequence of the formation of more isodimensional micelles — comparable with the opening action with a coacervate.

Fig. 44. Benzene is added to a 9.1% solution of Na oleate, after which the influence on the thickness of the micelle (A) and on the viscosity (B) is measured (HESS, KIESSIG, and PHILIPPOFF, 1941).

[1] P. KOETS and H. G. BUNGENBERG DE JONG, Protoplasma, 30 (1938) 206.
[2] From determinations of the increase of the long spacing on addition of various paraffins to a solution of K-laurate it can be calculated that a part of the paraffin is not taken up between the layers of the soap molecules. There are then two other possibilities: 1. a part of the paraffin is laid down parallel to the hydrocarbon chains in the lamellar micelles — which would agree with our idea, 2. a part of the paraffin would be taken up in the ionic micelles — which in our opinion will however hardly be present in a K-laurate solution of high concentration.
See E. W. HUGHES, W. M. SAWYER, and J. R. VINOGRAD, J. Chem. Phys., 13 (1945) 131.
[3] K. HESS, H. KIESSIG, and W. PHILIPPOFF, Fette und Seifen, 48 (1941) 377.

6°. It must be possible that in certain cases such an equilibrium is present between the forces A and B (Fig. 36) that a soap solution already exhibits elastic properties without addition of electrolyte, and thus contains large flat micelles. This case has already been mentioned (see p. 684).

7°. The discovery of SAMIS and HARTLEY[1] that the size of the micelles depends on the counter ions, also fits into our scheme well. The more strongly these ions are attached, the larger the micelle will be.

In our opinion these examples demonstrate the value of our way of looking at the problem. Emphasis must finally be laid on one thing. In spite of the fact that we consider the phenomena in soap solutions throughout as equilibrium phenomena, we use terms as "micelle", "coacervate", and so on, which on account of their colloid chemical past call forth ideas of strictly determined boundary surfaces (FREUNDLICH's Kapillarchemie). We wish however to retain these terms without crediting the boundary surface of micelle-equilibrium liquid with a separate significance. We thus look upon a micelle in a soap solution as a formation which is in equilibrium with the rest of the solution but which through its large dimensions and its structure has properties which the soap molecule as such does not possess. It is only with this restriction that we wish to continue to speak of micelles, coacervates, etc.

§ 6. ASSOCIATION OF ORGANIC DYES IN SOLUTION

With many organic dyes the same phenomena can be observed which occur also in soap solutions: a difference between the osmotic and the electrochemical properties. It is not our intention to explain completely all the properties of these dyes: it must suffice to remark that the problem of the colloidal solutions of organic dyes must be viewed in the same way as the question of the soap solutions. We may demonstrate this view with a single example.

Various polymethine dyes (pinacyanol, pseudoisocyanine) are soluble in alcohol and acetone as well as water. In the first two solvents the spectrum does not change as the concentration is raised; in these cases the dyes follow BEER's law. On the other hand if the solvent is water, important modifications take place in the spectrum (SCHEIBE et al[2]). These changes which are associated with a great rise of the viscosity, can best be explained by an association of the ions of the dye, first to double ions, later to polymolecular formations.

SCHEIBE here uses the word polymerisation, in our opinion incorrectly. Since we are here dealing with a completely reversible process, which is determined exclusively by the concentration of the substance and by the temperature, we must speak of association. There is no question that a macromolecule is formed; the final result is a micelle which is built up of many molecules (or ions) and which by its structure has different properties from the constituent parts.

To illustrate the changes in the spectrum on increase of the concentration the result of an experiment bij SCHEIBE[3] with pinacyanol may be added (Fig. 45). The conductivity of these solutions does not change so that the conclusion is drawn that there is association of the ions. In this medium — water — we must assume that the main spring in association are the VAN DER WAALS forces between the aromatic

[1] C. S. SAMIS and G. S. HARTLEY, Trans. Faraday Soc., 34 (1938) 1288.
[2] G. SCHEIBE, L. KANDLER, and H. ECKER, Naturwiss., 25 (1937) 75.
[3] G. SCHEIBE, Kolloid-Z., 82 (1938) 1.

hydrocarbon radicles. Association is also observed in a mixture of amyl alcohol and benzene. Here however the association is without influence on the spectrum and it is obvious that the nature of the association will be different; the polar groups will be the association centres. At first sight it will cause astonishment that the conductivity does not change with the association of these dye molecules. This happens because here the ratio cross section/charge does not change; not spherical but thread-like micelles are formed by association between the aromatic nuclei of the dye molecules.

Fig. 45.   Change in the spectrum of pinacyanol on increase of the concentration (SCHEIBE, 1938).

With pseudoisocyanine N-N' diethylchloride the association phenomena are still more remarkable. In this case also the spectrum changes in a similar way on increase of the concentration. The band, which must be attributed to the associated molecules, is very narrow and a fluorescence band occurs at precisely the same place. This band makes its appearance suddenly and the intensity rises rapidly. At the same concentration the viscosity increases considerably (a solution of $0.6 \cdot 10^{-2}$ mol exhibits practically no increase of the viscosity; a solution of 10^{-2} mol is a stiff gel), while a change in the conductivity also occurs. The various facts remind us strongly of the phenomena found in soap solutions. Streaming double refraction is also observed in concentrated solutions. These experiments lead to the idea that the molecules set themselves with the plane aromatic rings on each other as a result of which long threads can be produced. The association — it is not surprising — is strongly dependent on the temperature; rise in temperature makes the degree of association decrease.

The spectral bands which are produced by the association (briefly association bands) are not situated at the same position with different derivatives of pseudoisocyanine. If now we measure the association bands of a mixture of two dyes of this type, we find that these bands lie in between the bands of the component substances (SCHEIBE [1]). There is no question that the association bands of the mixture might be produced by an overlapping of the original bands. They are new bands and we must draw the conclusion from this that the electrons of the constituent molecules are no longer independent of each other. They must be amalgamated to an absorption and fluorescence unit.

The long association threads will be bent at random in solution. In a streaming

---

[1] loc. cit., p. 718.

solution they align themselves parallel to the direction of flow. Now it is important that only linear polarized light with the electric vector parallel to the direction of flow (thus in the length of the micelle) is strongly absorbed in the narrow association band. Light with the electric vector perpendicular to the direction of flow (thus in the plane of the molecules) is absorbed in the other two bands. The explanation of this must be as follows. The p-electrons responsible for the light absorption can (with the single molecules) only absorb light with the vector parallel to the plane of the aromatic rings. On association the molecules (from the optical point of view) amalgamate to a larger unit, as a result of which the absorption can only take place along the length of the micelles (thus perpendicular to the plane of the aromatic rings).

Various organic compounds are able to extinguish the fluorescence of the associated molecules. It is worth investigating this extinction with a few dyes (KATHEDER[1]). Let us first of all examine the fluorescence of quinoline red, a dye which is related to pseudoisocyanine. This dye has a normal fluorescence. It can be excited by every
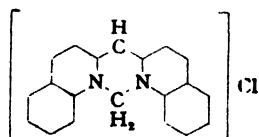
wavelength belonging to the absorption spectrum. The temperature has practically no influence on the fluorescence, while the latter rises gradually as the concentration increases. Addition of a foreign compound (pyrocatechol for example) can extinguish this fluoresence. Many molecules of pyrocatechol are then required for this purpose per molecule of dye. One only observes the first traces of extinction of the fluorescence when two to three molecules of pyrocatechol are present per molecule of quinoline red.

The behaviour of mono-naphtho-pseudoisocyanine N - N′ diethyl chloride is quite different. The fluorescence is very dependent on the temperature; a decrease of the fluorescence is associated with a decrease in the viscosity. The fluorescence starts suddenly at a certain concentration and rapidly reaches it maximum.
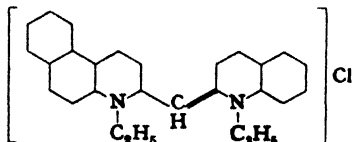
We already know that this fluorescence is connected with the formation of the micelles.

With these association colloids the fluorescence is strongly dependent on the pH. A shift of the pH from 7 to 8 makes the fluorescence (and the association) rise considerably. KATHEDER's explanation that in this case the hydration of the OH ions would extract water from the aggregate seems extremely improbable. Each OH ion (at pH = 8!) would have to bind a very large number of water molecules. It does in fact seem more probable that just as in the soaps a suppression of the dissociation will promote the association. The charge of the dye stuff ions will indeed be of

[1] F. KATHEDER, Kolloid-Z., 92 (1940) 299, 93, (1940) 28.

decisive influence. Addition of electrolyte in general must have precisely the same action; this is probably the cause of a lowering of the pH also having a favourable influence on the association. It can be predicted that in this direction different counter ions will have a different action.

Pyrocatechol also has a favourable action on the association (and thus on the intensity of the fluorescence) at low concentrations. KATHEDER again thinks of hydration of the pyrocatechol in this connection. We can safely assume that there will be hardly any question of this. We think rather of the correspondence with soap solutions where organic non-electrolytes will similarly have primarily a condensing action, thus an action favouring association. At higher concentrations the pyrocatechol begins to extinguish the fluorescence. Only one pyrocatechol molecule is however required for the purpose to many dye molecules. Thus with pseudoisocyanine one molecule of pyrocatechol to 10^3 to 10^6 dye molecules produces the effect. Not only this, but at higher concentrations of the dyes (that is, at the same time a higher degree of association) relatively less pyrocatechol is required to bring about a just observable decrease of the fluorescence. .

The idea of SCHEIBE et al[1] is that the foreign molecules are taken up at the end of the chain and that the light energy — which can be propagated over the aggregate without energy loss — is delivered to these molecules. In principle it must be possible for several quanta to be delivered to a single disturbing molecule in a short time. An idea which brings to mind the problem of photosynthesis. Would it not be possible that a similar mechanism is present also in this case? The coupling of the electrons in an aggregate of the colouring matters under discussion results in a quantum not remaining attached to the molecule which has absorbed this quantum. The quantum can be transported over large distances and emitted elsewhere in fluorescence or given up elsewhere by transfer of the energy to another molecule.

We again see agreement with the soap coacervates (of, for example, the oleate) in which other fatty acid ions (for example, the undecylate) show a strongly opening action. A single undecylate molecule can in like manner reduce the association of many oleate molecules. It would be very well worth while to look in the case of the pseudoisocyanines also for agreement and difference of disturbing molecule and substrate as was done in the above mentioned investigation on soap coacervates.

By adsorption on certain surfaces it is possible so to adsorb pseudoisocyanin from solutions, in which no association is possible, that the association band makes its appearance. This phenomenon occurs especially on mica, while glass and quartz are also suitable (see SKERLAK[2]). The previous considerations make it already possible for us to see what conditions such a surface must satisfy; naturally there must be negatively charged spots at which the positive dye ions will be attached. This adsorption does not succeed with polystyrene for example and this must be due to the non-polar character of this substance. Such an adsorption on mica apparently forces the molecules of the dye into a particular position so that the structure of the adsorbed layer becomes similar to that of the micelles. Not exactly so, since the spectral band is somewhat displaced on most adsorbents. SKERLAK's investigation on the adsorption of pseudo-isocyanin on mica gives a beautiful confirmation of this hypothesis. The molecules of the dye stand perpendicular to the adsorbing surface.

[1] G. SCHEIBE, A. SCHÖNTAG, and F. KATHEDER, Naturwiss., 27 (1939) 499.
[2] T. SKERLAK, Kolloid-Z., 95 (1941) 267.

We have the feeling that it will be important in the study of the colloid chemistry of the colouring matters to lay emphasis on the correspondence with soap solutions. The aligning forces exerted on the dye molecules in solution are to be compared with the forces to which soap molecules are subject. Both kinds of molecule are anisodimensional, a strongly polar group is attached to a large non-polar portion. It is not surprising that correspondences are found in the solutions[1].

"The biological importance of paracrystals in general can hardly be overestimated" writes BATEMAN (1945). We have already shed light on the problem of permeability where studies of oleate coacervates furnish the possibility of approximating certain sides of the problem. Although we do not wish to regard the soap micelles directly as paracrystals, the somewhat ordered structure in these formations is conspicuous nevertheless. There can be no doubt that correspondences between this ordering and the structure of protoplasm can be found here. Consequently the study of the association colloids is an important contribution towards many biological problems.

---

[1] See also E. RABINOWITCH and L. F. EPSTEIN, *J. Am. Chem. Soc.*, 63 (1941) 69.

AUTHOR INDEX

Page

Carothers, W. H. 605
Carter, W. C. and Scott, R. L. and Magat,
  M. 169
Cassie, A. B. D. 89, 522, 542, 548
Centola, G. 501, 546, 584
Clark, G. L, 611 (see Gross, S. J. and —),
  621, 623 (see Sisson, W. A. and —), 631
Clark, G. L. and Parker, E. A. and
  Schaade, J. A. and Warren, W. J. ... 616
Claude, G. 525
Clews, C. B. J. 611
Coelingh, M. B. ... 522, 523, 531, 532, 533
Cohan, L. H. 531
Cohn, E. J. and Ed;all, J. T. 186, 192
Cohn, E. J. and Green, A. A. and
  Blanchard, M. H. 191
Colmant, P., 555 (see Brönsted, N. and —)
Contoni 575
Coolidge, A. S. 528, 529
Coolidge, A. S. and Lamb, A. D. 527
Cotton and Mouton 510
Corey, R. B. and Wyckoff, R. W. G... 616
Corrin, M. L., 696 (see Harkins, W. D.
  and Mattoon, R. W. and —), 713 (see
  Harkins, W. D. and Mattoon, R. W.
  and —)
Couette 115
Cragg, L. H. 168, 175
Crowfoot, D. 240 (see Bernal, J. D. and —),
  616 (see Bernal, J. D. and —), 616
Curme, G. O. and Douglas, S. D. 664
Cuthbertson, A. C. and Gee, G. and
  Rideal, E. K. 43

D

Dale 576, 591
Danes, V. Z. 110
Debye, P. 87, 100, 104, 113, 146, 148, 149
Debye-Scherrer 614, 620
Dekker, W. A. L., 205 (see Bungenberg de
  Jong, H. G. and —), 229 (see Bungen-
  berg de Jong, H. G. and —), 322 (see
  Bungenberg de Jong, H. G. and —), 326
  (see Bungenberg de Jong, H. G. and —),
  341 (see Bungenberg de Jong, H. G.
  and —), 345 (see Bungenberg de Jong,
  H. G. and —), 360 (see Bungenberg de
  Jong, H. G. and —), 401 (see Bungen-
  berg de Jong, H. G. and — and Winkler,
  K. C.), 407 (see Bungenberg de Jong, H.
  G. and — Linde, P. van der), 415 (see
  Bungenberg de Jong, H. G. and Wink-
  ler, K. C.), 509 (see Bungenberg de Jong,
  H. G. and —)
Derksen, J. C. 499, 500, 501, 502 (see Katz,
  J. R. and — and Bon, F. W.), 502 (see
  Katz, J. R. and —), 502, 545 (see Katz,

Page

J. R. and —), 546 (see Katz, J. R. and —),
  548 (see Katz, J. R. and —), 559 (see
  Katz, J. R. and —) 559, 566
Derksen, J. C. and Katz, J. R. and Hess,
  K. and Trogus, C. 595
Desmaroux, J. and Mathieu, M. 121
Desmaroux, J. and Mathieu, M. and Petit-
  pas, T. 548
Devoto, G. 102
Dexter, S. T. 500 (see Kraemer, E. O. and
  —), 501, 577 (see Kraemer, E. O. and —)
Dhar, N. R., 573 (see Prahash, S. and —)
Dickenson, S., 616 (see Astbury, W. T.
  and — and Bailley, K.)
Diehl, J. M. and Iterson, G. van 595
Dieterichi, C. 576
Dinglinger, A. 132, 141 (see Schulz, G. V.
  and —). 144, 152 (see Schulz, G. V. and
  —), 169 (see Schulz, G. V. and —)
Dippel, C. J. 499 (see Boer, J. H. de and —)
Dobry, A. 135, 145, 238, 508, 580
Döhle, W., 545 (see Staudinger, H. and
  —), 582 (see Staudinger, H. and —)
Donnan, F. G. 560, 561, 562, 564
Donnan, F. G. and Krishnamurti, K... 500
Donnan, F. G. and Potts, H. E... 683, 712
Doormaal, P. M. van 183 (see Pfeiffer,
  J. Ph. and —), 657 (see Pfeiffer, J. Ph.
  and —)
Dostal, H., 41 (see Mark, H. and —), 54, 123
Doty, P. M., 146 (see Zimm, B. H. and
  Stein, R. S. and —)
Doty, P. M. and Zimm, B. H. and Mark,
  H. 146
Douglas, S. D., 664 (see Curme, G. O. and
  —)
Douglas, S. D. and Stoops, W. N...... 664
Duclaux, J. 84, 119, 135, 504, 511, 581
Duhem 56
Dürr, W., 45 (see Freudenberg, K. and —)

E

Ecker, H., 718 (see Scheibe, G. and
  Kandler, H. and —)
Eckersall, N., 536 (see Urquhart, A. R.
  and —)
Eckling, K. and Kratky, O. 641
Edelman, H. J. 203 (see Kruyt, H. R. and
  —)
Eder, K., 150 (see Staudinger, H. and —)
Edsall, J. T., 186 (see Cohn, E. J. and —),
  186 (see Anson, M. L. and —), 192 (see
  Cohn, E. J. and —), 240 (see Anson,
  M. L. and —)
Ehlers, G. 163
Eilers, H. 176, 178
Einstein, A. 166, 180

SUBJECT INDEX

Page